Stochastic volatility of volatility and variance risk premia

Ole E. Barndorff-Nielsen
The T.N. Thiele Centre for Mathematics in Natural Science & CREATES, Aarhus University
oebn@imf.au.dk

Almut E. D. Veraart
Imperial College London & CREATES
a.veraart@imperial.ac.uk

December 15, 2011

Abstract

This paper introduces a new class of stochastic volatility models which allows for stochastic volatility of volatility (SVV): *Volatility modulated* non–Gaussian *Ornstein–Uhlenbeck* (VMOU) processes. Various probabilistic properties of (integrated) VMOU processes are presented. Further we study the effect of the SVV on the leverage effect and on the presence of long memory.

One of the key results in the paper is that we can quantify the impact of the SVV on the (stochastic) dynamics of the variance risk premium (VRP). Moreover, provided the physical and the risk–neutral probability measures are related through a structure–preserving change of measure, we obtain an explicit formula for the VRP.

Keywords: Stochastic volatility of volatility, Lévy process, Ornstein–Uhlenbeck process, variance risk premium, supOU process;

JEL classification: C10, C13, C14, G10.

1 Introduction

Stochastic volatility (SV) is one of the key concepts in financial econometrics and has been studied extensively in recent years. Numerous empirical studies have revealed the fact that asset price volatility is time–varying and exhibits clusters, and a good volatility model is therefore essential in various applications such as portfolio selection, option pricing and risk management.

This paper contributes to the growing literature on *continuous–time* stochastic volatility models by introducing a new class of stochastic volatility models which allows for *stochastic volatility of volatility*. Here we view stochastic volatility of volatility as expressing the possibility or fact that there is greater variability i.e. more volatility in the data structure under study than might initially be surmised. In modelling terms this means that we consider the initial thinking as embodied in a (classical) SV model and want to describe the extra variability by a further source of randomness.

Clearly, there are many classical SV models which can serve as the base SV process in introducing stochastic volatility of volatility. Here we have decided in favour of the non–Gaussian Ornstein–Uhlenbeck (OU) process, which is driven by a Lévy subordinator and was introduced by Barndorff-Nielsen & Shephard (2001, 2002), for modelling squared volatility. The reason for this choice is that such models are appealing due to both their analytical tractability and their good empirical performance. The former is due to the fact that OU processes are defined by a linear stochastic differential equation, which can be solved explicitly. Also, the integrated OU process – which in the volatility context represents the accumulated stochastic variability over a certain time period – has an explicit representation in terms of the OU process itself and the driving Lévy process. The latter property has been revealed in various empirical studies: For instance, Barndorff-Nielsen & Shephard (2002) showed the good empirical performance of superpositions of OU models for modelling stochastic volatility in exchange rate data; Nicolato & Venardos (2003) focused on derivative pricing based on OU volatility models, and recently Benth (2011) showed their applicability in commodity markets.

Another interesting feature of Lévy–driven OU processes is that their extreme behaviour can be rather diverse, since it is determined by the tail behaviour of the increments of the driving Lévy process, see e.g. Fasen et al. (2006), Klüppelberg & Lindner (2010). Note also that recent empirical research by Christensen et al. (2011) based on ultra high frequent asset price data suggests that rapid price movements might be often caused by rapid movements in the volatility rather than by price jumps. This finding calls for new classes of stochastic volatility models which can account for such rapid moves.

In this paper, we use OU processes driven by a Lévy subordinator as the base model for stochastic volatility. In a second step, we introduce stochastic volatility of volatility by volatility modulation of the driving Lévy subordinator. This can be done in three ways: Integrating an additional volatility process with respect to the Lévy subordinator, or applying a stochastic time change to the subordinator, or combining these two forms of volatility modulation. Then one obtains a *volatility modulated Lévy* process. Such a volatility modulated Lévy subordinator is then used as the driving process of an OU process, which results in a *volatility modulated OU* (VMOU) process. The VMOU is then used as a new stochastic volatility process in an asset price model.

The focus in this paper on the new class of models based on VMOU processes reflects our viewpoint that a concrete full specification of what is meant by volatility of volatility has to be relative to a given or chosen basic volatility model so that volatility of volatility means variation beyond what is contained in the base model. Thus, for instance, the quadratic variation of the volatility in the base model is not volatility of volatility in the sense of that tenet.

1.1 Related literature

Before we study our new stochastic volatility model in more detail, we briefly set it into perspective to other stochastic volatility models in the recent literature.

The first generation of SV models was established with the focus on accounting for the well–know stylised facts of asset returns such as time varying volatility, volatility clusters, the existence of a leverage effect, see Nelson (1991), i.e. the (typically negative) correlation between asset returns and their volatility, see e.g. Barndorff-Nielsen & Shephard (2007, 2012), Ghysels et al. (1996), Shephard & Andersen (2009), Shephard (2005) for a review. In a next step, the classical stochastic volatility models were extended to allow for jumps, long memory, long run components and non–linear mean–reversion etc., see e.g. Comte & Renault (1998).

Also, the existence of implied volatility smiles and skews derived from option prices clearly indicated that SV is an essential component in an asset pricing model, see also Cox (1996), Dupire (1994), Hagan et al. (2002), Heston (1993), Stein & Stein (1991). In particular, the class of *multifactor* SV models is important to mention in this context. They form a very natural generalisation of the classical one factor SV models and are very successful in the context of option pricing, see e.g. Christoffersen et al. (2009, 2008).

Clearly, various additional random factors can be introduced in a stochastic volatility model in different ways. The classical multifactor SV models usually work with a linear combination of SV models. However, an extra source of randomness could also be added to reflect a *stochastic leverage* component, see Veraart & Veraart (2012).

Alternatively, we can work with a richer random structure in the volatility itself, the approach we pursue in this paper. This route has also been taken earlier by Meddahi & Renault (2004). They introduced a semiparametric class of volatility models which is characterized by an autoregressive dynamic of the stochastic variance, which is called *square–root stochastic autoregressive volatility* (SR-SARV) and which is shown to be closed under temporal aggregation.

It is clear that there are various ways to construct multifactor SV models (both in discrete and in continuous time) and the question of which approach is preferable will crucially depend on the specific application.

Our new class of VMOU processes can be viewed as a specific multifactor SV model, where the additional source of randomness enters as the volatility of the driving process of the SV. The interesting feature of our particular choice of a multifactor SV model is that it can be linked explicitly to the so–called *variance risk premium* (VRP), which has recently attracted a lot of research attention. Recall that an investor faces at least two sources of uncertainty, when investing in a security: the uncertainty about the return (which is described by the return variance) and the uncertainty about the return variance itself, see Carr & Wu (2009). It turns out that the risk associated with the uncertainty in the variance is measured by the VRP and recent empirical work indicates that the VRP exhibits stochastic dynamics itself, see for instance Carr & Wu (2009), Bollerslev et al. (2009), Todorov (2010) and Drechsler & Yaron (2011). This finding raises the question of how such stochastic dynamics of the VRP can be modelled. We will later give one possible answer to this important question.

1.2 Key results and outline

Next we list the main contributions of the paper, before we describe them in more detail in the following sections.

First, we find that our new volatility model is highly analytically tractable: We derive its cumulant function and second order structure explicitly, and we get a representation result for the integrated squared volatility process. The latter result reveals that the long term behaviour of integrated volatility in our new modelling framework exhibits volatility clusters itself. Further, it transpires that the additional volatility of volatility component has an important impact on both the *memory*, i.e. on the autocorrelation structure, of the volatility process and on the possibility of incorporating the leverage effect into the asset price model, see Section 2 for more details.

Next, we discuss changes of measure to a risk–neutral probability measure (in an incomplete market) and focus in particular on so–called structure–preserving measure changes. We then explain, how option prices can be computed in our new modelling framework based on Fourier inversion techniques, see Section 3.

Section 4 contains the important result which links the stochastic volatility of volatility component to the dynamics of the *variance risk premium*. More precisely, in the special case where the risk–neutral and the physical probability measures are linked by a structure–preserving change of measure, we can show that the stochastic volatility of volatility solely determines the stochastic dynamics of the variance risk premium. The fact that the stochastic volatility of volatility drives the variance risk premium has been demonstrated in the context of an equilibrium model based on economic theory by Bollerslev et al. (2009) and Drechsler & Yaron (2011). However, it is interesting to see that we can confirm this result (under suitable assumptions) based on a purely probabilistic model.

In this paper, we moreover show that the additional volatility of volatility can actually be used to add additional memory to a SV model. An alternative method for allowing for both long memory and additional stochastic volatility of volatility simultaneously is presented in Section 5, where we discuss an extension of our new modelling framework to the class of *volatility modulated supOU processes*.

Section 6 concludes and gives an outlook on future research. The proofs of our main results are relegated to the Appendix.

2 The new modelling framework

Throughout the paper, we assume that the logarithmic asset price $Y = (Y_t)_{t\geq 0}$ is given by an Itô semimartingale

$$dY_t = a_t dt + \sigma_{t-} dW_t + dJ_t, \tag{1}$$

which is defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, where $a = (a_t)_{t\geq 0}$ is a predictable drift process, $\sigma = (\sigma_t)_{t\geq 0}$ is a càdlàg stochastic volatility process and $J = (J_t)_{t\geq 0}$ is the pure jump component of the Itô semimartingale. Note that an Itô semimartingale is defined as a semimartingale whose characteristics are absolutely continuous with respect to the Lebesgue measure (see e.g. Jacod (2008)).

The variation of financial markets, which is often referred to as squared *volatility*, is usually measured by means of the quadratic variation of the logarithmic price process. In our modelling framework, the quadratic variation (QV) (denoted by $[\cdot]$) is given by

$$[Y]_t = \sigma_t^{2+} + \sum_{0\leq s\leq t} (\Delta J_s)^2, \tag{2}$$

where $\sigma_t^{2+} = \int_0^t \sigma_s^2 ds$ is the integrated squared stochastic volatility process and where $\Delta J_s = J_s - J_{s-}$ denotes the jump of J at time s. Taking the square root of the quadratic variation $\sqrt{[Y]_t}$ leads to a measure of the *volatility* of the asset price. In the following, we will work with a specific new model for the squared volatility process which is given by a volatility modulated non–Gaussian Ornstein–Uhlenbeck process.

2.1 The volatility modulated non–Gaussian Ornstein–Uhlenbeck process

Barndorff-Nielsen & Shephard (2001, 2002) proposed to model the squared volatility σ^2 by a non–Gaussian Ornstein–Uhlenbeck (OU) process. In the following, we will refer to such a model as *BNS model*. Here, we generalise such OU processes to allow for an additional *stochastic volatility of volatility component*. The new class of processes is called *volatility modulated* non–Gaussian *Ornstein–Uhlenbeck* (VMOU) processes. They are defined as follows.

For $t \geq 0$, let $\sigma_t^2 := V_t^{\upsilon,\tau} := V_t$, where

$$dV_t = -\lambda V_t dt + dL_{\lambda t}^{\upsilon,\tau}, \tag{3}$$

where $\lambda > 0$ is a constant, the *memory parameter*, and $L^{\upsilon,\tau}$ is the background driving volatility modulated Lévy process given by

$$dL_{\lambda t}^{\upsilon,\tau} = \upsilon_{\lambda t-} dL_{\tau_{\lambda t}}, \qquad \text{where} \qquad \tau_t = \int_0^t \xi_{s-} ds, \tag{4}$$

and where $(L_t)_{t\geq 0}$ is a Lévy subordinator with characteristic triplet $(\gamma, 0, \nu)$, i.e.

$$\mathbb{E}(\exp(i\theta L_t)) = \exp(t\psi_L(\theta)), \qquad \text{where} \qquad \psi_L(\theta) = i\theta\gamma + \int_0^\infty \left(e^{i\theta x} - 1\right)\nu(dx),$$

where $\gamma \geq 0$ and ν is a Lévy measure on $(0,\infty)$ satisfying $\int_0^\infty (x \wedge 1)\nu(dx) < \infty$ and also $\int_0^\infty (\log(x) \vee 1)\nu(dx) < \infty$.

Further $(\upsilon_t)_{t\geq 0}$ and $(\xi_t)_{t\geq 0}$ denote stationary positive, càdlàg stochastic (volatility) processes. In addition, we assume that L, υ and τ are mutually independent.

Note that we restrict ourselves to time change processes τ, which are absolutely continuous. Another popular class of time change processes is the class of Lévy subordinators. However, it is well–known that a Lévy subordinator time–changed by an independent Lévy subordinator is itself a Lévy subordinator, see e.g. Sato (1999, Theorem 30.4). Therefore this case is already included in our modelling framework.

The processes υ^2 and τ can be interpreted as the *stochastic variability of variance*. In the following, we will often also refer to υ, τ, ξ as *stochastic volatility of volatility* to simplify the exposition. Clearly, when $\upsilon_t \equiv 1$ and $\tau_t = t$, we obtain the well–known BNS model. Both υ and τ can be driven by a Brownian motion or (and) a jump process. E.g. we can think of υ being a Gamma–OU or IG–OU process, see Barndorff-Nielsen & Shephard (2001) or a square root diffusion, see Cox et al. (1985).

Note that the volatility modulated Lévy process $L^{\upsilon,\tau}$ is in fact volatility modulated in two ways: We have a stochastic integrand υ which scales the jump size of the subordinator by a stochastic factor, and we have a time change process τ which determines the speed at which the jumps occur. While stochastic proportional and temporal scaling are, under suitable regularity assumptions, equivalent in a Brownian framework and for stable Lévy processes, see Veraart & Winkel (2010), this is however not the case for general Lévy processes and, in particular, not for a general Lévy subordinator – the case we study here. In a concrete empirical application, it might well be sufficient to focus on one source of stochastic volatility of volatility only, but in our theoretical investigations, we wish to focus on the more general case.

Remark It is important to note that there are at least four different ways to include an additional stochastic component in a non–Gaussian OU process, two of which have been presented above. Alternatively, one could make the memory parameter λ stochastic and could study models of the form $dV_t = -\lambda_t V_t dt + dL_t$, where λ_t is a positive stationary process. Closely related to the latter is the concept of supOU processes and the need to model (quasi) long range dependence, see Barndorff-Nielsen & Shephard (2002), and these two ideas could be combined in a single construction. We come back to this case in Section 5.

Still another possibility would be to let $dV_t = V_{t-}dU_t + dL_t$, where (U, L) is a bivariate Lévy process, see e.g. Behme et al. (2011). However, this case and the case of a stochastic memory parameter λ_t appear less appealing in the present modelling context since we do not get an explicit formulae for the integrated process $V_t^+ = \int_0^t V_s ds$, which is regarded as a key quantity in financial econometrics. Such a representation is possible for our model defined in (3) and (4) and will be described in more detail below.

It turns out that asset price models where the logarithmic asset price is given by an Itô semimartingale and the squared stochastic volatility process is given by a VMOU process generally do not belong to the class of affine models as introduced by Duffie et al. (2003), see also Kallsen (2006) for a survey. However, if υ is constant and the density of the time change ξ is affine, then we obtain an affine model. Let us study a concrete example of such an affine representation next.

Example A stochastic volatility model which accounts for stochastic volatility of volatility and the leverage effect, i.e. the (possibly negative) correlation between the asset price and the volatility, could be defined by

$$\begin{aligned} dY_t &= a_t dt + \sqrt{V_{t-}} dW_t + \rho dL_{\tau_{\lambda t}}, \\ dV_t &= -\lambda V_t dt + \upsilon_{\lambda t-} dL_{\tau_{\lambda t}}, \\ \upsilon_t &\equiv 1, \\ d\tau_t &= \xi_{t-} dt, \\ d\xi_t &= \alpha(\beta - \xi_t)dt + \gamma\sqrt{\xi_t} dB_t, \end{aligned}$$

where $\rho \leq 0$, B denotes a Brownian motion (independent of W), $\alpha, \beta, \gamma > 0$ denote positive constants and the other quantities are defined as before. Note that this model belongs to the class of *affine* models since the the BNS model itself is affine and the new (additional) time change is given by the time integral of an affine process, see Keller-Ressel (2008).

In the following, we will study the key properties of our new modelling framework.

2.2 Properties of the volatility modulated Lévy process

Let us briefly discuss the main properties of the volatility modulated Lévy process $L^{\upsilon,\tau}$, which is the driving process of our new class of VMOU processes.

In the following, we will denote by Leb the Lebesgue measure.

2.2.1 Stochastic proportional

First, we focus on the process

$$L_t^{\upsilon,Id} = \int_0^t \upsilon_{s-} dL_s,$$

where $Id(\cdot)$ denotes the identity function and υ and L are independent. Then the characteristic function for a constant $\theta \in \mathbb{R}$ is given by

$$\mathbb{E}\left[\exp\left(i\theta L_t^{\upsilon,Id}\right)\right] = \mathbb{E}\left[\mathbb{E}\left[\exp\left(i\theta L_t^{\upsilon,Id}\right)\Big|\,\upsilon\right]\right] = \mathbb{E}\left[\exp\left(\int_0^t \psi_L(\theta\upsilon_{s-})ds\right)\right].$$

Also, the characteristic triplet of the semimartingale $L^{\upsilon,Id}$ is given by $\left(A(L^{\upsilon,Id}), 0, Leb \otimes \nu(L^{\upsilon,Id})\right)$, where

$$\begin{aligned} A(L^{\upsilon,Id})_t &= \gamma \int_0^t \upsilon_s ds, \\ \nu(L^{\upsilon,Id})_t(G) &= \int \mathbb{I}_G(\upsilon_{t-}x)\nu(dx), \end{aligned}$$

for any $G \in \mathcal{B}$ with $0 \notin G$, see Kallsen & Shiraev (2002, Lemma 3).

2.2.2 Stochastic time change

Next, we study the process

$$L_t^{1,\tau} = L_{\tau_t} = L_{\int_0^t \xi_{s-} ds},$$

where τ and L are independent.

Recall that we call a stochastic process X adapted w.r.t. τ (or $(\tau_t)_{t\geq 0}$–adapted) if X is constant on any interval $[\tau_{t-}, \tau_t]$ for any $t \geq 0$. Note that the time change satisfies $\tau_0 = 0$ and τ is continuous. In that case, L is τ–adapted, see e.g. Jacod (1979). As soon as we have adaptedness w.r.t. the time change, then important properties of the base process carry over to the time–changed process. In particular, we get

$$\mathbb{E}\left[\exp\left(i\theta L_t^{1,\tau}\right)\right] = \mathbb{E}\left[\mathbb{E}\left[\exp\left(i\theta L_t^{1,\tau}\right)\Big|\, \tau\right]\right] = \mathbb{E}\left[\exp\left(\psi_L(\theta)\tau_t\right)\right],$$

for a constant $\theta \in \mathbb{R}$, and L_τ has characteristic triplet $(\gamma\tau, 0, \tau \otimes \nu)$. Also, the *differential* characteristic triplet (w.r.t. the Lebesgue measure) is given by $(\gamma\xi_\cdot, 0, \xi_\cdot \otimes \nu))$, see Kallsen & Shiraev (2002, Lemma 5) and Barndorff-Nielsen & Shiryaev (2010, Theorem 8.4).

2.2.3 Combined volatility modulation

For the doubly volatility modulated Lévy process, we get from the above results (using the independence of L, υ and τ) that

$$\begin{aligned}\mathbb{E}\left[\exp\left(i\theta L_t^{\upsilon,\tau}\right)\right] &= \mathbb{E}\left[\mathbb{E}\left[\exp\left(i\theta L_t^{1,\tau}\right)\Big|\, \upsilon, \tau\right]\right] = \mathbb{E}\left[\exp\left(\int_0^t \psi_L(\theta\upsilon_{s-})d\tau_s\right)\right] \\ &= \mathbb{E}\left[\exp\left(\int_0^t \psi_L(\theta\upsilon_{s-})\xi_{s-}ds\right)\right].\end{aligned}$$

Further, $L^{\upsilon,\tau}$ has characteristic triplet $(A(L^{\upsilon,\tau}), 0, \tau \otimes \nu(L^{\upsilon,\tau}))$, where

$$\begin{aligned}A(L^{\upsilon,\tau})_t &= A(L^{\upsilon,Id})_{\tau_t} = \gamma\int_0^{\tau_t} \upsilon_s ds, \\ \nu(L^{\upsilon,\tau})_t(G) &= \nu(L^{\upsilon,Id})_{\tau_t}(G) = \int \mathbb{I}_G(\upsilon_{\tau_{t-}}x)\nu(dx),\end{aligned}$$

for any $G \in \mathcal{B}$ with $0 \notin G$.

2.3 Properties of the VMOU process

We have defined the VMOU process as the solution to the stochastic differential equation given by (3). From standard arguments, we can deduce the following representation:

$$V_t = V_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} dL_{\lambda s}^{\upsilon,\tau}.$$

Then, the stationary version of V can be written as

$$V_t = \int_{-\infty}^t e^{-\lambda(t-s)} dL_{\lambda s}^{\upsilon,\tau},$$

where $L^{\upsilon,\tau}$ is suitably extended to the negative half line (see Barndorff-Nielsen & Shephard (2001)). Note that a VMOU processes can be regarded as a special case of a Lévy semistationary ($\mathcal{LSS}$) process, which has recently been introduced by Barndorff-Nielsen et al. (2010).

Remark The model specification of the stationary VMOU processes has the property that the marginal distribution of V is independent of the parameter λ. Otherwise put, indicating the dependence on λ by writing $V(\lambda)$, i.e.

$$V_t(\lambda) := V_t = \int_{-\infty}^{t} e^{-\lambda(t-s)} dL_{\lambda s}^{\upsilon,\tau} = e^{-\lambda t} \int_{-\infty}^{t} e^{\lambda s} dL_{\lambda s}^{\upsilon,\tau} = e^{-\lambda t} \int_{-\infty}^{\lambda t} e^{u} dL_{u}^{\upsilon,\tau} = V_{\lambda t}(1).$$

So we have seen that the law of $V_t(\lambda)$ (which by the stationarity does not depend on t) equals the law of $V_{\lambda t}(1)$. Due to the stationarity, the law of $V_{\lambda t}(1)$ equals the law of $V_0(1)$, which implies that it does not depend on λ. Since the parameter λ has no impact on the marginal distribution of V, but only determines the autocorrelation structure, see Proposition 2 below, we can interpret λ as the *memory parameter*.

Clearly, the Lévy subordinator L is a Markov process. However, the Markov property is not preserved under stochastic integration or general stochastic time change. In particular, V is no longer a Markov process. However, the bivariate process (V, υ) satisfies the Markov property if υ is itself a Markov process and $\tau_t = ct$ for a constant $c > 0$.

Due to the high analytical tractability of our new model the characteristic function of V conditional on υ and τ can be directly computed (hence we omit the proof).

Proposition 1 *The conditional characteristic function of the VMOU is given by*

$$\begin{aligned}\psi_{V_t}^{\upsilon;\tau}(\theta) &= \mathbb{E}(\exp(i\theta V_t)|\, \upsilon, \tau) = \mathbb{E}\left(\exp\left(i\theta \int_{-\infty}^{t} e^{-\lambda(t-s)} \upsilon_{\lambda s} dL_{\tau_{\lambda s}}\right)\middle|\, \upsilon, \tau\right) \\ &= \exp\left(\int_{-\infty}^{t} \psi_L(\theta e^{-\lambda(t-s)} \upsilon_{\lambda s-}) d\tau_{\lambda s}\right) = \exp\left(\int_{-\infty}^{t} \psi_L(\theta e^{-\lambda(t-s)} \upsilon_{\lambda s-}) \xi_{\lambda s-} \lambda ds\right).\end{aligned}$$

Also the second order structure of V can be easily derived. Throughout the paper, we will use the following notation. For $i \in \mathbb{N}$, we denote the ith cumulant of the Lévy subordinator L_t by $\kappa_i(L_t)$ (provided it exists). Clearly, we have $\kappa_i(L_t) = t\kappa_i(L_1)$ and, in particular, we write $\kappa_i := \frac{1}{\lambda}\kappa_i(L_\lambda) = \kappa_i(L_1)$. Also, we will write $\gamma(h) = Cov(\upsilon_{t-}\xi_{t-}, \upsilon_{(t+h)-}\xi_{(t+h)-})$ for $h > 0$.

Proposition 2 *The mean, variance and autocovariance of the stationary process V are given by*

$$\begin{aligned}\mathbb{E}(V_t) &= \kappa_1 \mathbb{E}(\upsilon_0)\mathbb{E}(\xi_0), \\ Var(V_t) &= \frac{1}{2}\kappa_2 \mathbb{E}\left(\upsilon_0^2\right)\mathbb{E}(\xi_0) + \kappa_1^2 \int_{-\infty}^{0}\int_{-\infty}^{0} e^x e^y \gamma(|x-y|) dx dy, \\ Cov(V_t, V_{t+h}) &= e^{-\lambda h}\frac{1}{2}\kappa_2 \mathbb{E}\left(\upsilon_0^2\right)\mathbb{E}(\xi_0) + e^{-\lambda h}\kappa_1^2 \int_{-\infty}^{0}\int_{-\infty}^{\lambda h} e^x e^y \gamma(|x-y|) dx dy.\end{aligned}$$

Proof The proof is given in Section A.1 in the Appendix. □

Recall that simple OU processes as well as the often used CIR process (Cox et al. (1985), Heston (1993)) have an exponentially declining autocorrelation function and, hence, do not allow for longer memory in the volatility process. However, many empirical studies reveal that medium or long memory is an important property of stochastic volatility and should be accounted for by a realistic model. From the autocorrelation structure we have derived in Proposition 2 we see clearly that the stochastic volatility of volatility components indeed generate a more slowly decaying autocorrelation function compared to a simple OU process.

As already indicated, an alternative method to introduce long(er) memory in an OU process is the concept of superposition studied by Barndorff-Nielsen & Shephard (2002). In fact, this concept can also be extended to the framework of a VMOU processes and will be discussed in detail in Section 5.

Finally, we state a representation result for the increments of a VMOU process.

Proposition 3 *For any $h \geq 0$, we have*

$$V_{t+h} - V_t = \left(e^{-\lambda h} - 1\right) V_t + \int_t^{t+h} e^{-\lambda(t+h-s)} dL^{\upsilon,\tau}_{\lambda s}. \tag{5}$$

Proof The proof is straightforward and hence omitted. □

2.3.1 Visualisation of some key model properties

In this section we aim to visualise the main features of the VMOU process by a brief simulation study.

Let us focus on the case where the time change is chosen to be the identity function, i.e. $\tau_{\lambda t} = Id(\lambda t) = \lambda t$. Hence the volatility modulation only appears through stochastic proportional υ, which is chosen to be a non–Gaussian OU process itself. More precisely, we simulate the following three processes:

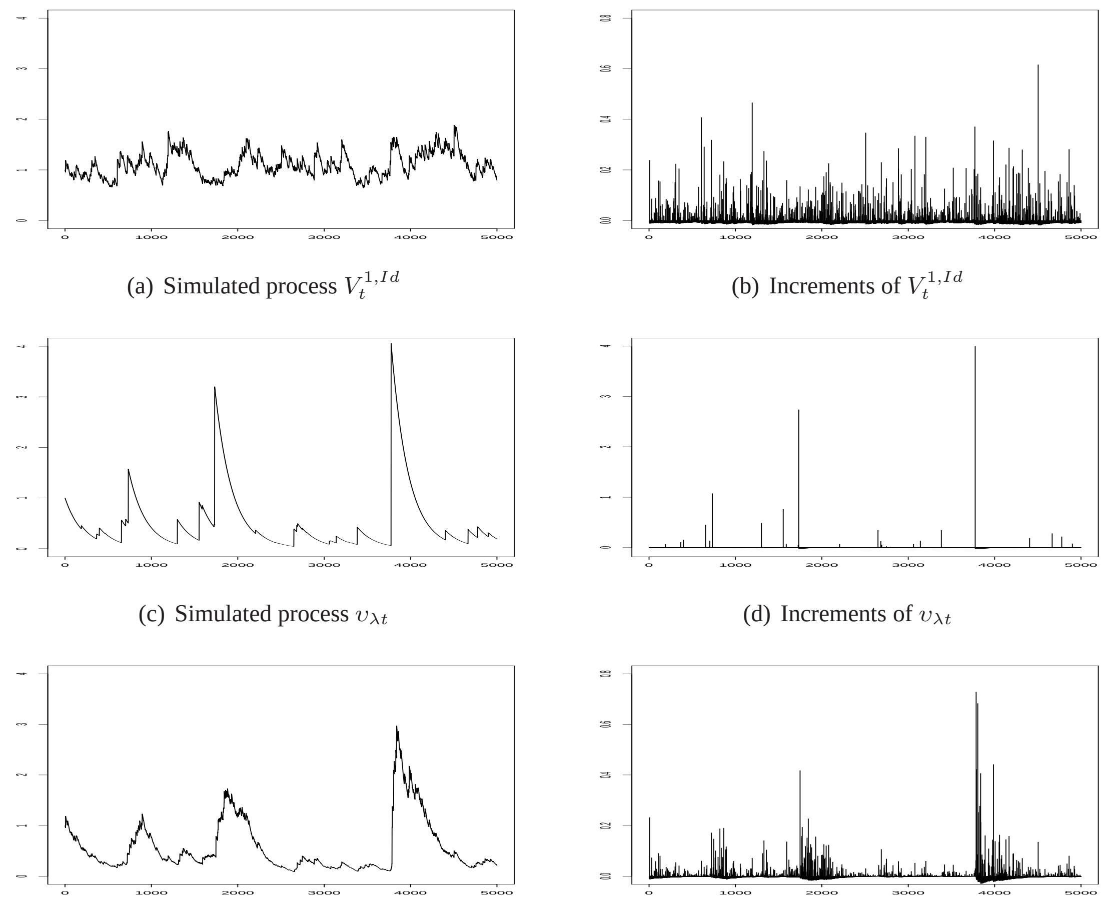
$$\begin{aligned} dV_t^{1,Id} &= -\lambda V_t dt + dL_{\lambda t}, \\ d\upsilon^2_{\lambda t} &= -\lambda^{(\upsilon)} \lambda \upsilon^2_{\lambda t} dt + dL^{(\upsilon)}_{\lambda^{(\upsilon)}\lambda t}, \\ dV_t^{\upsilon,Id} &= -\lambda V_t dt + \upsilon_{\lambda t-} dL_{\lambda t}. \end{aligned}$$

We simulate 5000 observations with a step size of 1 using the Euler scheme. Note that the driving Lévy subordinators L and L^{υ} are chosen to be two independent Gamma processes. Recall that the $\Gamma(a, s)$–density function is given by $f(x) = \frac{1}{s^a \Gamma(a)} x^{a-1} e^{-x/s}$, which implies that the corresponding random variable has a mean of as and a variance of as^2. Here we specify $\lambda = 0.01$ and $L_1 \sim \Gamma(a, s)$ for $a = 10$ and $s = 0.1$. Further, we choose $\lambda^{(\upsilon)} = 1$ and $L_1^{(\upsilon)} \sim \Gamma(a_2, s_2)$ with $a_2 = 0.1$ and $s_2 = 1/a_2$. Note that this parameter choice ensures that $V^{1,Id}$ and $V^{\upsilon,Id}$ have the same mean.

The simulated paths of the three processes and their corresponding returns are depicted in Figure 1. We can clearly spot the various theoretical properties we have just discussed: First, the volatility of volatility introduces additional volatility in the sense that we obtain a process, the VMOU $V^{\upsilon,Id}$, which has volatility clusters itself. This fact becomes rather apparent when comparing the plots of the returns of the two processes $V^{1,Id}$ and υ with the returns of the VMOU process $V^{\upsilon,Id}$. Note also, that bursts in volatility can build up more gradually in the new modelling framework, whereas in a standard BNS model we obtain a sudden upwards jump achieving a local maximum immediately, followed by an exponential decay.

In Figure 2 we can see how the presence of a stochastic volatility component in an OU process can increase the memory in the sense that the autocorrelation function decreases more slowly than for a simple OU process.

The case of $V^{1,\tau}$, where the additional volatility enters through a stochastic time change does not require further visualisation, since it easy to imagine how volatility clusters can be obtained by changing the “speed” of the OU process stochastically.



(a) Simulated process $V_t^{1,Id}$

(b) Increments of $V_t^{1,Id}$

(c) Simulated process $\upsilon_{\lambda t}$

(d) Increments of $\upsilon_{\lambda t}$

(e) Simulated process $V_t^{\upsilon,Id}$

(f) Increments of $V_t^{\upsilon,Id}$

Figure 1: Simulation study: (a) & (b) Simulation of the OU process $V^{1,Id}$ and its increments; (c) & (d) Simulation of the volatility of volatility process $\upsilon_{\lambda t}$ and its increments; (e) & (f) Simulation of the VMOU process $V^{\upsilon,Id}$ and its increments.

2.4 Properties of the integrated VMOU process

Next we study the *integrated* stochastic volatility (IV). IV is regarded as a key object of interest in financial econometrics since it reflects the accumulated (continuous) quadratic variation over a certain period of time (usually a day). So, this section analyses main properties of this key quantity in our new modelling framework. In the following, we will use the notation $V^+ = (V_t^+)_{t\geq 0}$ for the integrated process

$$V_t^+ = \int_0^t V_s ds.$$

Also, we define

$$\epsilon_\lambda(t) := \frac{1}{\lambda}\left(1 - e^{-\lambda t}\right).$$

First of all, we derive a representation result for the integrated process.

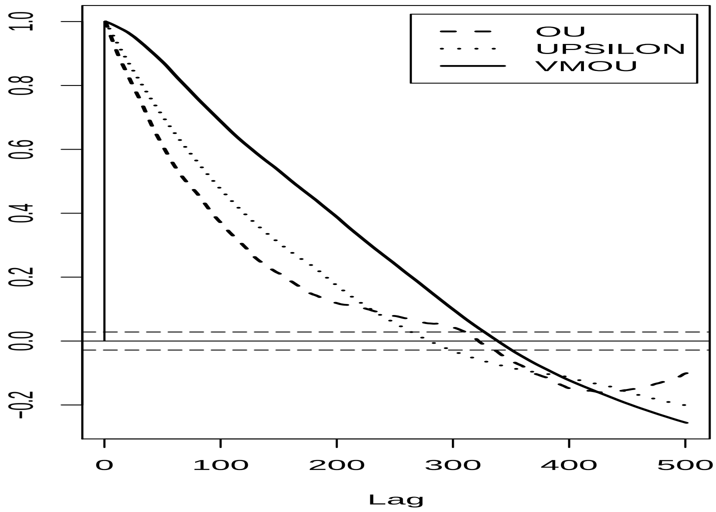


Figure 2: Autocorrelation functions of $V_t^{1,Id}$, $\upsilon_{\lambda t}$ and $V_t^{\upsilon,Id}$.

Proposition 4 *The integrated process can be written as*

$$V_t^+ = \epsilon_\lambda(t)V_0 + \int_0^t \epsilon_\lambda(t-s)dL_{\lambda s}^{\upsilon,\tau} = \frac{1}{\lambda}\left(L_{\lambda t}^{\upsilon,\tau} + V_0 - V_t\right).$$

The proof of the above Proposition is straightforward and, therefore, not given here.

These different representations of V^+ are interesting, since they shed some light on the joint behaviour of V and V^+. In particular, we can deduce some results on co–jumps and cointegration (as introduced by Granger (1981)). Clearly, V_t and $L_{\lambda t}^{\upsilon,\tau}$ have identical jumps (breaks), they co–break, i.e. $\Delta V_t = \Delta L_{\lambda t}^{\upsilon,\tau}$, but V and $L_{\lambda t}^{\upsilon,\tau}$ are not cointegrated. However, V^+ and $L_{\lambda t}^{\upsilon,\tau}$ are in fact cointegrated since

$$\lambda V_t^+ - L_{\lambda t}^{\upsilon,\tau} = V_0 - V_t.$$

I.e. we have found a linear combination of the non–stationary processes V^+ and $L_\lambda^{\upsilon,\tau}$ which is stationary. So, roughly, for large t, λV_t^+ will have the same distribution as $L_{\lambda t}^{\upsilon,\tau}$, where the error in this approximation is a stationary process, which is given by $V_0 - V_t$. Now we can clearly see which influence the stochastic volatility of volatility has in the new modelling set up: While the long–run behaviour of integrated volatility in the classical BNS model is described by the background driving Lévy process, our new model allows for a greater flexibility in the sense that it can allow for processes which have stationary, but not necessarily independent increments in the long run behaviour of the integrated variance. In particular, the long–run behaviour of integrated volatility can exhibit volatility clusters itself due to the new component of volatility of volatility. This is clearly an important aspect of the additional volatility of volatility component.

Also, since $L_\lambda^{\upsilon,\tau}$ is a nonnegative process, the integrated process V^+ is bounded below by the quantity $\epsilon_\lambda(t)V_0$.

Note that we can use formula (5) to derive a representation result for the increments of the integrated process: For $h \geq 0$ we get

$$V^+_{t+h} - V^+_t = \epsilon_\lambda(h)V_t + \int_t^{t+h} \epsilon_\lambda(t+h-s)dL^{\upsilon,\tau}_{\lambda s}. \quad (6)$$

The various cumulants of the integrated process can now be easily derived using the representation result from Proposition 4.

2.5 Some comments on model identification

When introducing an additional volatility component in an OU model, we have to ask ourselves whether such models are in fact identifiable. This question can be addressed via the characteristic functional of the VMOU processes. Note in particular that in order to ensure that our model is uniquely identifiable, we have to impose parameter restrictions on L, υ, and τ. It might be difficult to distinguish between the two sources of stochastic volatility of volatility. In order to shorten the exposition we present the results for the case when both sources of additional volatility are present. In a concrete application, however, it might be sufficient to set either υ or the density process ξ to one and work with one source of stochastic volatility of volatility only.

Further, can we test from real data that an additional volatility component is present in the data? In Section 3 we will answer this question by linking the volatility of volatility component to the variance risk premium. However, without using any risk–neutral information available to us through option prices, how can we distinguish between a non–Gaussian OU process and a volatility modulated non–Gaussian OU process statistically?

One way to answer this question is to focus on the quadratic variation of the VMOU process, which is given by

$$[V]_t = \int_0^t \upsilon^2_{\lambda s-} d[L]_{\tau_{\lambda s}}.$$

We know that for a Lévy subordinator L, $[L]$ is again a Lévy subordinator.

In the case that $\tau_t = t$, we see that $[V]$ has independent increments if υ is deterministic. As soon as υ is a stochastic process, the independent increment property generally does not hold any longer. Also, if υ was a deterministic function which is not just a constant, then $[V]$ does not have stationary increments any longer. A practical implication of these results is that, in principle, we can estimate the quadratic variation of the spot volatility process V (based on a spot volatility estimator, see e.g. Aït-Sahalia & Jacod (2009), Bandi & Renò (2008), Kristensen (2010), Lee & Mykland (2008), Veraart (2010)) and test statistically whether the estimated $[V]$ has independent and stationary increments.

In case of $\upsilon \equiv 1$, we have $[V]_t = [L]_{\tau_{\lambda t}}$, which is generally not a Lévy subordinator anymore since the independent increment property is violated. As before, this property may be tested statistically based on an empirical estimate of $[V]$, which is constructed based on a spot volatility estimator.

2.6 Leverage through stochastic volatility of volatility

Next, we show that the additional stochastic volatility of volatility component can be used for introducing the leverage effect into stochastic volatility models in a novel way. The (usually negative) correlation between asset returns and volatility has been found in many empirical studies, see e.g. Black (1976), Christie (1982) and Nelson (1991) among others and, more recently, by Harvey & Shephard (1996), Bouchaud et al. (2001), Tauchen (2004, 2005), Yu (2005) and Bollerslev et al. (2006).

So far, leverage type effects have usually been introduced by directly correlating the driving process of the volatility with the driving process of the asset prices (as e.g. in the Heston (1993) model). Introducing leverage in the BNS model is slightly more complicated since the volatility is driven by a subordinator and the price is driven by a Brownian motion which are inherently independent from each other (by the Lévy – Khintchine formula). Hence Barndorff-Nielsen & Shephard (2001) suggested to add a jump component to the asset price, which is given by the subordinator which drives the volatility multiplied by a (negative) constant. Such a structure assumes linear dependence between asset price and volatility.

However, having an additional random factor in the stochastic volatility model, i.e. the stochastic volatility of volatility, makes it possible to introduce leverage type effects indirectly and independently of the fact whether we want to have a jump component in the model for the logarithmic asset price. In order to illustrate this, let us look at a small example.

For simplicity, we discard jumps in the price process in the following and show that a leverage effect can be solely introduced by a diffusion component.

Example We consider the following model

$$\begin{aligned} dP_t &= \sqrt{V_{t-}} dW_t, \\ dV_t &= -\lambda V_t dt + \upsilon_{\lambda t-} dL_{\tau_{\lambda t}}, \\ \tau_t &= t, \\ d\upsilon_t &= \alpha(\beta - \upsilon_t)dt + \gamma\sqrt{\upsilon_t} B_t, \end{aligned}$$

for parameters $\lambda, \alpha, \beta, \gamma > 0$ and a Brownian motion $B = (B_t)_{t\geq 0}$ with $d[B, W]_t = \widetilde{\rho} dt$, for $\widetilde{\rho} \in [-1, 1] \setminus \{0\}$ and all the other quantities are defined as above. For this model, the correlation of the price and the volatility and the third moment of the price can be expressed in terms of the model parameters and, in particular, we get

$$Cov(P_t, V_t) = \mathbb{E}(P_t V_t) \not\equiv 0 \qquad \text{and} \qquad Cov(P_t, P_t^2) = \mathbb{E}\left(P_t^3\right) \not\equiv 0.$$

Note that the proof of the result is given in Section A.2 of the Appendix.

So, we see that we can have a non–zero correlation between the asset price and the squared volatility, even if the volatility is jump driven and there are no jumps in the logarithmic asset price.

Similarly, when we study the time change case, we can introduce a leverage effect in the diffusion part by correlating the Brownian motion driving the asset price with a Brownian motion driving the time change.

3 Change of measure and option pricing

So far, we have studied various model properties under the physical probability measure. Next, we discuss how a change of measure to a *risk–neutral* probability measure can be carried out, and we will furthermore describe how option prices can be computed in our new modelling framework.

3.1 Risk–neutral probability measures in incomplete markets

Recall that we denote by $\mathbb{P}$ the physical probability measure on the filtered probability space given by $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$. Let Z denote a positive martingale with mean 1. Then we define a new (equivalent) probability measure for $t, h \geq 0$ by

$$\mathbb{Q}_t(A) := \mathbb{E}(\mathbb{I}_A Z_t)) = \int_A Z_t d\mathbb{P}, \qquad \text{for all } A \in \mathcal{F}_t.$$

Clearly, $\mathbb{Q}_t$ is a probability measure on $\mathcal{F}_t$ with

$$\frac{d\mathbb{Q}_t}{d\mathbb{P}} = Z_t. \tag{7}$$

Furthermore, since Z is a martingale, we have $\mathbb{Q}_{t+h}(A) = \mathbb{Q}_t(A)$ for all $A \in \mathcal{F}_t$.

Recall that we call a probability measure $\mathbb{Q}$ which is equivalent to $\mathbb{P}$ the *risk–neutral probability measure* if all discounted price processes are $\mathbb{Q}$–martingales. In complete markets, such an equivalent martingale measure is unique. However, in this paper we focus on an *incomplete market* and, hence in order to do arbitrage–free option pricing, we need to find *a* risk–neutral probability measure, which is generally not unique. Thus the arbitrage–free option prices depend on the choice of the probability measure. This implies that different choices of the risk–neutral probability measure typically result in different option prices. Or more generally, we get a range of option prices depending on the choice of the risk–neutral probability measure. In the following, we will denote by $\mathcal{Q}$ the *set of all risk–neutral probability measures* for our modelling framework. So which risk–neutral measure should one choose? In the literature, some common classes of martingale measures are often used, such as martingale measures obtained from the Esscher transform, the minimal martingale measures, the minimum entropy martingale measures and the class of structure–preserving martingale measures. Such measure changes have been studied in much detail by Hubalek & Sgarra (2009) in the context of the BNS model.

3.1.1 Structure–preserving change of measure

Let us focus on the class of structure–preserving changes of measure in more detail. Suppose the stochastic processes describing the asset price, the stochastic volatility and the stochastic volatility of volatility components follow a set of stochastic differential equations (SDE). We call a change of measure *structure–preserving*, if also under the new probability measure we have the same structure of the SDEs. That means that if a SDE was driven by a Brownian motion or a Lévy processes, this property is preserved under the change of measure, but the corresponding model parameters typically change, and also the distribution of the Lévy process might change. Furthermore independence properties between stochastic processes carry over in the case of a structure–preserving change of measure.

Nicolato & Venardos (2003) derived a complete characterisation of the class of structure–preserving changes of measures for the BNS model. Also Veraart & Veraart (2012) studied structure–preserving changes of measure for the (generalised) Heston model. Since our modelling framework is a direct generalisation of the BNS set up, and the volatility of volatility components enter as independent factors, we can obtain structure–preserving changes of measure for our modelling framework as soon as the stochastic volatility of volatility components υ^2 and ξ follow either a non–Gaussian OU process or a square–root diffusion process, by combining the results in Nicolato & Venardos (2003) and Veraart & Veraart (2012). More precisely, one concrete structure–preserving change of measure can be obtained by constructing the product measure of a structure–preserving change of measure for the BNS model with a structure–preserving change of measure for the stochastic volatility of volatility components.

3.2 Option pricing

A popular method for computing option prices is based on the Laplace or Fourier transform of the asset price. Such methods have been introduced by Heston (1993) and have subsequently been studied by Carr & Madan (1999), Lee (2004), Lewis (2001), Raible (2000) amongst others.

Let us briefly describe the main idea of this approach as e.g. reviewed by Nicolato & Venardos (2003). Suppose we would like to compute the price of a European option with payoff function $c(Y_T)$ at time of maturity $T = t + h$ for $h > 0$. Recall that the stock price is given by $S_t = S_0 \exp(Y_t)$. I.e. for a European call, we have $c(Y_T) = (S_0 \exp(Y_T) - K)^+ = (S_T - K)^+$, where K denotes the strike price. Then the option price at time $t \leq T = t + h$ is given by

$$C_t = \mathbb{E}^{\mathbb{Q}} \left[e^{-r(T-t)} c(Y_T) | \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}} \left[e^{-rh} c(Y_{t+h}) | \mathcal{F}_t \right],$$

where $r \geq 0$ denotes the risk–free interest rate. Now, let $\widehat{c}$ denotes the Laplace transform of the payoff function c of an option, i.e.

$$\widehat{c}(z) = \int_{-\infty}^{\infty} e^{-zx} c(x) dx,$$

see Raible (2000) for explicit forms. If the Laplace transform of the logarithmic asset price, which we denote by ϕ, is known, then the option price C_t can be computed using Fourier inversion, specifically

$$C_t = \frac{e^{-rh}}{2\pi i} \int_{\varphi - i\infty}^{\varphi + i\infty} \phi(z) \widehat{c}(z) dz,$$

where φ is a constant belonging to the set where both $\widehat{c}$ and ϕ are defined (provided such a constant exists).

Here, we will show that a transform–based method can be used for computing option prices when the logarithmic asset price is given by the generalised BNS model where the squared volatility process is a volatility modulated non–Gaussian OU process. In order to do that, we derive the Laplace transform of the integrated volatility process and of the log–price process, which are both obtained in semi–analytic form.

Recall that we need to choose a risk–neutral probability measure $\mathbb{Q}$. Here we work with a *structure–preserving* change of measure. To simplify the exposition, we hence assume that the model is directly specified under the risk–neutral probability measure. Further, throughout this section, we will work with the following more specific model for the asset price which is given by $S_t = S_0 \exp(Y_t)$ for

$$dY_t = (\mu + \beta V_t) dt + \sqrt{V_{t-}} dW_t, \tag{8}$$

where $\mu, \beta \in \mathbb{R}$ and V are defined as before (but now we assume that the definition of the model holds under the risk–neutral measure $\mathbb{Q}$).

Before we derive the Laplace transformation of the price process, we formulate a condition which ensures that the discounted asset price $e^{-rt} S_t$ (for $r > 0$) is a (local) martingale, where we follow closely Nicolato & Venardos (2003). Applying Itô's formula, we obtain the following dynamics for the asset price S:

$$dS_t = S_{t-} \left(b_t dt + \sqrt{V_{t-}} dW_t \right), \qquad \text{where} \qquad b_t = \mu + \left(\beta + \frac{1}{2} \right) V_t.$$

Hence, the discounted asset price is a (local) martingale if and only if

$$b_t - r = 0. \tag{9}$$

Assuming the martingale condition holds, the Laplace transformation of the price process (provided it exists) can be determined in the form given in the following Proposition, the proof of which is given in the Appendix.

Proposition 5 *Let $\phi(\theta) = \mathbb{E}_t^{\mathbb{Q}}(\exp(\theta Y_{t+h}))$, for $h \geq 0$. Then*

$$\begin{aligned}\phi(\theta) &= \exp\left(\theta Y_t + \theta\mu h + \left(\beta\theta + \frac{\theta^2}{2}\right)\epsilon_\lambda(h)V_t\right)\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\int_t^{t+h}\chi_L(f(s,\theta))d\tau(\lambda s)\right)\right] \\ &= \exp\left(\theta Y_t + \theta\mu h + \left(\beta\theta + \frac{\theta^2}{2}\right)\epsilon_\lambda(h)V_t\right)\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\lambda\int_t^{t+h}\chi_L(f(s,\theta))\xi_{\lambda s-}ds\right)\right],\end{aligned}$$

where $f(s,\theta) := (\beta\theta + \theta^2/2)\epsilon_\lambda(t+h-s)\upsilon_{\lambda s-}$ and where χ_L denotes the log–transformed Laplace transform of L.

Note that in the case when υ and τ are deterministic, then we have an analytic formula for the Laplace transform of the conditional distribution of Y. In the stochastic case, the integral has to be evaluated using Monte Carlo methods. Finally, we can use the Fourier inversion formula presented above for computing the option price based on our new stochastic volatility model.

4 Variance risk premia

So far, we have seen that the additional stochastic volatility of volatility component can be motivated both from an empirical point of view when studying asset price data under the physical measure, since the additional component introduces more flexibility in the behaviour of the integrated variance, and also when studying option prices since the integrated variance also enters directly in the option pricing formula.

In this section we will show that the additional stochastic volatility of volatility component also plays a key role in determining the dynamics of the *variance risk premium* (VRP). We will demonstrate that, under a structure–preserving change of measure, the stochastic dynamics of the VRP are determined solely by the volatility of volatility component.

4.1 General result

In order to understand the influence of the volatility of volatility term even better, we study the *variance risk premium* (VRP), which has been studied extensively in the very recent literature, see e.g. Carr & Wu (2009), Bollerslev et al. (2009), Todorov (2010), Drechsler & Yaron (2011), Wu (2011).

Recall that the variance risk–premium is defined as the wedge between the conditional expectation of the quadratic variation over a future period in time under the physical and under the risk–neutral probability measure, see e.g. Todorov (2010) for more details. Therefore the standardised variance risk premium over the interval $[t, t+h]$ is given by

$$VRP_{t,t+h}^{\mathbb{Q}} := \frac{1}{h}\left[\mathbb{E}_t([Y]_{[t,t+h]}) - \mathbb{E}_t^{\mathbb{Q}}([Y]_{[t,t+h]})\right],$$

where $\mathbb{E}_t(\cdot) := \mathbb{E}(\cdot|\mathcal{F}_t)$ and $\mathbb{Q} \in \mathcal{Q}$ is a risk–neutral probability measure. Note that the VRP depends on the particular risk–neutral probability measure $\mathbb{Q}$.

Since, we have an explicit formula for the integrated squared volatility process, we can specify a fairly explicit general formula for the variance risk premium in the following Proposition, which is proved in the Appendix.

Proposition 6 *(i) Let $\mathbb{Q} \in \mathcal{Q}$. The variance risk premium is given by*

$$\begin{aligned} VRP_{t,t+h}^{\mathbb{Q}} &= \frac{1}{h}\left[\mathbb{E}_t\left(\zeta(t,t+h)\right) - \mathbb{E}_t^{\mathbb{Q}}\left(\zeta(t,t+h)\right)\right] \\ &= \frac{1}{h}\mathbb{E}_t\left(\zeta(t,t+h)\left(1-\frac{Z_{t+h}}{Z_t}\right)\right), \end{aligned} \tag{10}$$

where

$$\zeta(t,t+h) := \int_t^{t+h} \epsilon_\lambda(t+h-s)dL_s^{\upsilon,\tau} + \sum_{t\le s\le t+h}(\Delta J_s)^2,$$

and Z is the martingale associated with $\mathbb{Q}$ through (7).

(ii) The range of the variance risk premium is given by

$$\{VRP_{t,t+h}^{\mathbb{Q}}, \mathbb{Q} \in \mathcal{Q}\}.$$

Formula (10) contains a fairly explicit formula for the variance risk premium associated with the stochastic volatility model given by a volatility modulated non–Gaussian OU process. In particular, we see that the volatility of volatility component influences the dynamics of the variance risk premium in quite a direct fashion through the process $L^{\upsilon,\tau}$.

Also, note that there are two parts of the variance risk premium: One is due to the continuous martingale part of the log–price and one is due to the jump part. In the following, we will denote by $VRP_{t,t+h}^{\mathbb{Q},c}$ the part of the VRP due to the continuous component in the price and by $VRP_{t,t+h}^{\mathbb{Q},d}$ the part due to the jumps J, i.e.

$$\begin{aligned} VRP_{t,t+h}^{\mathbb{Q},c} &= \frac{1}{h}\left[\mathbb{E}_t\left(\zeta^c(t,t+h)\right) - \mathbb{E}_t^{\mathbb{Q}}\left(\zeta^c(t,t+h)\right)\right], \\ VRP_{t,t+h}^{\mathbb{Q},d} &= \frac{1}{h}\left[\mathbb{E}_t\left(\zeta^d(t,t+h)\right) - \mathbb{E}_t^{\mathbb{Q}}\left(\zeta^d(t,t+h)\right)\right], \end{aligned}$$

where

$$\zeta^c(t,t+h) := \int_t^{t+h} \epsilon_\lambda(t+h-s)dL_s^{\upsilon,\tau}, \qquad \zeta^d(t,t+h) := \sum_{t\le s\le t+h}(\Delta J_s)^2.$$

Corollary 7 *The continuous and the jump part of the variance risk premium are given by*

$$\begin{aligned} VRP_{t,t+h}^{\mathbb{Q},c} &= \frac{1}{h}\mathbb{E}_t\left[\left(\int_t^{t+h} \epsilon_\lambda(t+h-s)dL_s^{\upsilon,\tau}\right)\left(1-\frac{Z_{t+h}}{Z_t}\right)\right], \\ VRP_{t,t+h}^{\mathbb{Q},d} &= \frac{1}{h}\mathbb{E}_t\left[\left(\sum_{t\le s\le t+h}(\Delta J_s)^2\right)\left(1-\frac{Z_{t+h}}{Z_t}\right)\right]. \end{aligned}$$

From this general result, we see that there are five stochastic processes which impact the variance risk premium: The Lévy subordinator L, the two SVV components υ and τ, the jump process in the logarithmic asset price J and the martingale Z. In the next section, we will formulate conditions which will allow us to quantify their influence even more explicitly.

4.2 Variance risk premia under structure–preserving changes of measures

It turns out that we can obtain an even more explicit result for the VRP if we add the following two assumptions.

(A1) The jump process in the asset price, denoted by J, is a pure jump Lévy process.

(A2) The risk–neutral probability measure $\mathbb{Q} \in \mathcal{Q}$ is obtained by a structure–preserving change of measure.

Proposition 8 *Let $\mathbb{Q} \in \mathcal{Q}$. Suppose J is a pure jump Lévy process under both $\mathbb{P}$ and $\mathbb{Q}$. Then the part of the VRP due to jumps is given by a constant, in particular,*

$$VRP_{t,t+h}^{\mathbb{Q},d} = \int_{\mathbb{R}} x^2 \nu_J(dx) - \int_{\mathbb{R}} x^2 \nu_J^{\mathbb{Q}}(dx),$$

where $\nu_J(dx)dt$ and $\nu_J^{\mathbb{Q}}(dx)dt$ are the predictable compensators of the Poisson random measure associated with J under $\mathbb{P}$ and under $\mathbb{Q}$, respectively.

Note that Proposition 8 clearly holds under the stronger assumptions (A1) and (A2).

Next, let us study the the VRP due to the continuous part of the price process.

Proposition 9 *Assume that assumption (A2) holds. Let $\eta_t^{\upsilon,\tau} := \upsilon_{\lambda t-}\xi_{\lambda t-}$. Then we have*

$$VRP_{t,t+h}^{\mathbb{Q},c} = \frac{1}{h}\Big[\eta_t^{\upsilon,\tau}\left(\mathbb{E}(L_1) - \mathbb{E}^{\mathbb{Q}}(L_1)\right)(h - \epsilon_\lambda(h)) + \lambda \int_t^{t+h} \epsilon_\lambda(t+h-s)\left(\mathbb{E}(L_1)\mathbb{E}_t\left(\eta_s^{\upsilon,\tau} - \eta_t^{\upsilon,\tau}\right) - \mathbb{E}^{\mathbb{Q}}(L_1)\mathbb{E}_t^{\mathbb{Q}}\left(\eta_s^{\upsilon,\tau} - \eta_t^{\upsilon,\tau}\right)\right) ds\Big]. \tag{11}$$

Note that L is a subordinator and, hence, $\mathbb{E}(L_1) > 0$. Also, under the structure–preserving change of measure, the predictable compensator of L changes and, hence, $\mathbb{E}(L_1) - \mathbb{E}^{\mathbb{Q}}(L_1) \neq 0$.

Remark We see that the stochastic proportional υ and the density of the time change ξ play the same role in determining the dynamics of the variance risk premium.

The above propositions show that, given a structure–preserving change of measure, the stochastic dynamics of the variance risk premium are solely determined by the stochastic volatility of volatility component υ and τ, respectively. If these terms were not stochastic, then the variance risk premium would be deterministic, which would contradict recent empirical findings e.g. by Drechsler & Yaron (2011), Bollerslev et al. (2009).

The formula (11) in Proposition 9 can be computed explicitly under an additional modelling assumption:

Corollary 10 *Assume that assumption (A2) holds. Consider the following two cases: Assume that $\eta_t^{\upsilon,\tau} = \eta_t^{(i)}$, where $\eta_t^{(1)} := \upsilon_{\lambda t-}$ (i.e. $\xi \equiv 1$) and $\eta_t^{(2)} := \xi_{\lambda t-}$ (i.e. $\upsilon \equiv 1$). Also, for $i = 1, 2$, let*

$$d\eta_t^{(i)} = a\left(b - \eta_t^{(i)}\right)dt + g\sqrt{\eta_t^{(i)}}dB_t^{\eta},$$

where a, b, g are positive constants satisfying the Feller condition $2ab > g^2$ (and $\upsilon_0^{(i)} > 0$) and B^{η} is a standard Brownian motion.

Then we get for $i = 1, 2$ that

$$VRP^{\mathbb{Q},(i)c}_{t,t+h} = \eta^{(i)}_t F^{(1)}(h) + F^{(2)}(h),$$

where $F^{(1)}(h)$ and $F^{(2)}(h)$ are explicitly known deterministic functions, given in the Appendix.

Also in the case of a non–Gaussian OU process, we get explicit results, see the following Corollary.

Corollary 11 *Assume that assumption (A2) holds. Consider the following two cases again: Assume that $\eta^{\upsilon,\tau}_t = \eta^{(i)}_t$, where $\eta^{(1)}_t := \upsilon_{\lambda t-}$ (i.e. $\xi \equiv 1$) and $\eta^{(2)}_t := \xi_{\lambda t-}$ (i.e. $\upsilon \equiv 1$). Also, for $i = 1, 2$, let*

$$d\eta^{(i)}_t = -a\eta^{(i)}_t dt + dL^{\eta}_{at},$$

where $a > 0$ and L^{η} is a Lévy subordinator. Then we get for $i = 1, 2$ that

$$VRP^{\mathbb{Q},(i)c}_{t,t+h} = \eta^{(i)}_t G^{(1)}(h) + G^{(2)}(h),$$

where $G^{(1)}(h)$ and $G^{(2)}(h)$ are explicitly known deterministic functions, given in the Appendix.

So, we have obtained an explicit formula for the VRP which depends on the physical and risk–neutral parameters of the underlying model. Further, we see that the stochastic dynamics of $VRP_{t,t+h}$ (as a stochastic process in t with fixed $h > 0$) are determined by the volatility of volatility component $\eta^{\upsilon,\tau}$.

Note that the classical BNS model (assuming a structure–preserving change of measure) implies that the variance risk premium is deterministic. This fact can be easily seen from our results above. However, by including a stochastic volatility of volatility term, we allow for *stochastic dynamics of the variance risk premium*.

4.2.1 Comments on the results

Note that both Bollerslev et al. (2009) and Drechsler & Yaron (2011) also established a link between a volatility of volatility component and the VRP. However, they studied a self–contained general equilibrium model and show that the variance risk premium is solely driven by the volatility of (consumption growth) volatility, where this explicit formula has been derived using a log–linear approximation.

Here, we do not work with an equilibrium model, but extend one of the popular (probabilistic) asset price models, the BNS model, to allow for an additional volatility of volatility component. In order to derive our explicit formula for VRP and to establish the link to the volatility of volatility component, we did not need any approximation, due to the high analytic tractability of our new model. However, the main assumption we made in this Subsection was that the physical and the risk–neutral probability measures are related through a structure–preserving change of measure, which seems to be a strong, but nevertheless rather natural assumption from a modelling point of view. Let us elaborate on the latter aspect in more detail.

As we have already mentioned earlier, incomplete markets have the property that the risk–neutral probability measure is not uniquely determined, leading to a range of risk–neutral probability measures. Consequently, the variance risk premium is not uniquely determined, since it depends on the choice of the risk–neutral probability measure.

Recent empirical studies indicate that the dynamics of the variance risk premium are stochastic. How can we explain such stochastic dynamics?

One possible answer is given in this paper: If we work with the VMOU stochastic volatility model and a structure–preserving change of measure, we find that the SVV drives the VRP and, in particular, that we obtain stochastic dynamics of the VRP.

Another possibility would be to work with a classical SV model which does not allow for SVV and at the same time apply a more sophisticated change of measure, which potentially could also induce stochastic dynamics of the VRP.

5 Introducing long memory: Volatility modulated supOU processes

Finally, we will give an outlook on further extensions of our new modelling framework. Here we show how long(er) memory can be incorporated into the class of VMOU processes. This is done by extensions of the idea of supOU processes, as introduced by Barndorff-Nielsen (2001) and further discussed in Barndorff-Nielsen & Shephard (2003), Barndorff-Nielsen & Leonenko (2005), Fasen & Klüppelberg (2007), Barndorff-Nielsen & Stelzer (2011), Barndorff-Nielsen & Stelzer (2012).

The long(er) memory can be introduced by randomising the memory parameter λ using the concept of *Lévy bases*. A Lévy basis is an independently scattered random measure whose values are infinitely divisible. The foundation of the theory of such measures were laid by Rajput & Rosinski (1989), see also Pedersen (2003). For a recent account of the definition and basic properties of Lévy bases see Section 1.3 of Barndorff-Nielsen et al. (2011*a*) and also Barndorff-Nielsen (2011), Barndorff-Nielsen & Shephard (2012) for further reviews.

5.1 Background on Lévy bases

Throughout the paper, we denote by $\mathcal{B}$ the family of Borel sets in $\mathbb{R}^k$ for $k \in \mathbb{N}$ and by $\mathcal{B}_b$ the subfamily of bounded elements of $\mathcal{B}$.

Definition 12 *Let $M = \{M(B) : B \in \mathcal{B}_b\}$ be a collection of random variables on some probability space $(\Omega, \mathcal{A}, P)$. We call M an* independently scattered random measure *(ISRM) if, for every sequence $\{B_n\}$ of mutually disjoint sets in $\mathcal{B}_b$ where $\cup_{n=1}^{\infty} B_n \in \mathcal{B}_b$, the random variables $M(B_n)$ are independent for $n = 1, 2, \ldots$, and also $M(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} M(B_n)$ almost surely.*

We are particularly interested in ISRM which are infinitely divisible.

Definition 13 *Let L be an ISRM on $\mathbb{R}^k$. We call L a* Lévy basis *if for all $B \in \mathcal{B}_b$ the random variable $L(B)$ is infinitely divisible.*

Rajput & Rosinski (1989) have shown that every Lévy basis has a Lévy–Khintchine representation of the form

$$\begin{aligned} C\{\theta \ddagger L(B)\} &= \log\left(\mathbb{E}(\exp(i\theta L(B)))\right) \\ &= i\theta a(B) - \frac{1}{2}\theta^2 b(B) + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - i\theta x \mathbb{I}_{[-1,1]}(x)\right) n(dx, B), \end{aligned} \tag{12}$$

where a is a signed measure on $\mathbb{R}$, b is a measure on $\mathbb{R}$, $n(\cdot, \cdot)$ is the generalised Lévy measure such that $n(dx, B)$ is a Lévy measure on $\mathbb{R}$ for fixed B and a measure on $\mathbb{R}^k$ for fixed dx. Further, along the lines of Rajput & Rosinski (1989) we define the so–called *control measure*.

Definition 14 *Let L be a Lévy basis with Lévy–Khintchine representation (12). The* control measure *c is then defined by*

$$c(B) = |a|(B) + b(B) + \int_{\mathbb{R}} \min(1, x^2) n(dx, B), \tag{13}$$

where $|\cdot|$ denotes the total variation.

Next we define the Radon–Nikodym derivatives of the three components of c, when we differentiate with respect to c. We have

$$\alpha(s) = \frac{da}{dc}(s), \qquad \beta(s) = \frac{db}{dc}(s), \qquad \nu(dx, s) = \frac{n(dx, \cdot)}{dc}(s). \tag{14}$$

In particular, we obtain $n(dx, ds) = \nu(dx, s)c(ds)$. In the following, we assume without loss of generality that $\nu(dx, s)$ is a Lévy measure for each fixed s.

Definition 15 *We call $(\alpha, \beta, \nu(dx, \cdot), c) = (\alpha(s), \beta(s), \nu(dx, s), c(ds))_{s \in \mathbb{R}^k}$ a* characteristic quadruplet *(CQ) associated with a Lévy basis L on $\mathbb{R}^k$ provided the following conditions hold:*

(i) Both α and β are measurable functions on $\mathbb{R}^k$, where β is restricted to be nonnegative.

(ii) For fixed s, $\nu(dx, s)$ is a Lévy measure on $\mathbb{R}$, and for fixed dx it is a measurable function on $\mathbb{R}^k$.

(iii) The element c is a measure on $(\mathbb{R}^k, \mathcal{B}_b)$ such that $\int_B \alpha(s)c(ds)$ is a (possibly signed) measure on $(\mathbb{R}^k, \mathcal{B}_b)$ and $\int_B \nu(dx, s)c(ds)$ is a Lévy measure on $\mathbb{R}$ for fixed $B \in \mathcal{B}$.

Altogether, one can show that every Lévy basis on $\mathbb{R}^k$ determines a CQ of the form $(\alpha, \beta, \nu(dx, \cdot), c) = (\alpha(s), \beta(s), \nu(dx, s), c(ds))_{s \in \mathbb{R}^k}$. And, conversely, every CQ satisfying the conditions in Definition 15 determines, in law, a Lévy basis on $\mathbb{R}^k$.

Now we get the following result for the cumulant function of the Lévy basis (presented in infinitesimal form).

$$\begin{aligned} C\{\theta \ddagger L(ds)\} &= \log\left(\mathbb{E}(\exp(i\theta L(ds))\right) \\ &= i\theta a(ds) - \frac{1}{2}\theta^2 b(ds) + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - i\theta x \mathbb{I}_{[-1,1]}(x)\right) n(dx, ds) \\ &= \left(i\theta\alpha(s) - \frac{1}{2}\theta^2\beta(s) + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - i\theta x \mathbb{I}_{[-1,1]}(x)\right) \nu(dx, s)\right) c(ds) \\ &= C\{\theta \ddagger L'(s)\}c(ds), \end{aligned} \tag{15}$$

where $L'(s)$ denotes the *Levy seed* of L at s. The Lévy seed is in fact an infinitely divisible random variable with Lévy–Khintchine representation

$$C\{\theta \ddagger L'(s)\} = i\theta\alpha(s) - \frac{1}{2}\theta^2\beta(s) + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - i\theta x \mathbb{I}_{[-1,1]}(x)\right) \nu(dx, s).$$

Note that one can associate a Lévy process with any Lévy seed.

In applications, we often work with special subclasses of Lévy bases, as defined in the following.

Definition 16 *Let L denote a Lévy basis on $\mathbb{R}^k$ with CQ given by $(\alpha, \beta, \nu(dx, \cdot), c)$.*

(i) If $\nu(dr, s)$ does not depend on s, we call L factorisable.

(ii) If L is factorisable and if c is proportional to the Lebesgue measure and $\alpha(s)$ and $\beta(s)$ do not depend on s, then L is called homogeneous.

Note that for a homogeneous Lévy basis, the associated Lévy seed $L'(s)$ does not depend on s.

5.2 Integrals with respect to Lévy bases

Integration with respect to Lévy bases can be done according to Rajput & Rosinski (1989), where integrals of deterministic kernel functions f with respect to a Lévy basis L were defined. In particular, we get from Rajput & Rosinski (1989, Proposition 2.6) that

$$C\{\theta \ddagger f \bullet L\} = \log\left(\mathbb{E}\left(\exp\left(i\theta \int f dL\right)\right)\right) = \int C\{\theta f(s) \ddagger L'(s)\} c(ds), \tag{16}$$

for a deterministic function f which is integrable with respect to the Lévy basis.

In order to allow for stochastic integrands (independent of the driving Lévy basis), we can think of a suitable extension (using conditioning) of the integration concept developed by Rajput & Rosinski (1989). Alternatively, one could employ the concept used by Walsh (1986), which is also described in detail in Barndorff-Nielsen et al. (2011*a*,*b*).

5.3 Introducing long memory through integrals w.r.t. Lévy bases

Now we show how long memory can be introduced in an Ornstein–Uhlenbeck process by randomising the memory parameter through the concept of a Lévy basis. First we review the basic supOU process and then we extend this process to allow for additional stochastic volatility of volatility.

5.3.1 SupOU model

Recall that the supOU process, as introduced by Barndorff-Nielsen (2001), is defined by

$$\widetilde{V}_t = \int_0^\infty \int_{-\infty}^t e^{-\lambda(t-s)} \tilde{L}(ds, d\lambda), \tag{17}$$

for a Lévy basis $\tilde{L}$ on $\mathbb{R} \times \mathbb{R}_+$ which has characteristic quadruplet

$$(0, 0, \nu(dx), dt\lambda\pi(d\lambda)). \tag{18}$$

Here ν is the Lévy measure of the Lévy subordinator, π denotes a probability measure on $(0, \infty)$ and $dt\lambda\pi(d\lambda)$ (with $(t, \lambda) \in \mathbb{R} \times \mathbb{R}_+$) is the control measure of $\widetilde{L}$, see Barndorff-Nielsen (2010) for more details. The supOU process defined in (17) can be regarded as an extension of an OU process driven by a Lévy subordinator. Here we associate a distribution with the memory parameter λ through the Lévy basis $\tilde{L}$. See Barndorff-Nielsen & Stelzer (2012) for conditions on the existence of the above integral.

We can easily compute the autocorrelation function (assuming square integrability of the Lévy basis $\widetilde{L}$). Let $\kappa_2 = \int_0^\infty x^2 \nu(dx)$. Then

$$\begin{aligned}
&\mathbb{E}\left((\tilde{V}_0 - \mathbb{E}(\tilde{V}_0))(\tilde{V}_t - \mathbb{E}(\tilde{V}_t))\right) = \kappa_2 \int_0^\infty \int_{-\infty}^0 e^{-\lambda t + 2\lambda s} ds \lambda \pi(d\lambda) \\
&= \kappa_2 \int_0^\infty e^{-\lambda t} \pi(d\lambda) \int_{-\infty}^0 e^{2\lambda s} \lambda ds = \int_0^\infty e^{-\lambda t} \pi(d\lambda) \underbrace{\kappa_2 \int_{-\infty}^0 e^{2u} du}_{:= X^{(1)}} \\
&= X^{(1)} \int_0^\infty e^{-\lambda t} \pi(d\lambda) = X^{(1)} \widehat{\pi}(\lambda),
\end{aligned}$$

where $X^{(1)}$ does not depend on λ and where $\widehat{\pi}(\lambda)$ denotes the Laplace transform of π.

Let us study an example, which shows how long memory can be obtained in the context of supOU processes, see Barndorff-Nielsen (2001), Barndorff-Nielsen & Shephard (2001) and Barndorff-Nielsen & Shephard (2012).

Example Let us assume that π is the Gamma law $\Gamma(2\overline{H}, 1)$ for $\overline{H} > 0$. The corresponding density function is given by

$$f_{\Gamma(2\overline{H},1)}(x) = \frac{1}{\Gamma(2\overline{H})} x^{2\overline{H}-1} e^{-x}, \qquad \text{for } x > 0.$$

Then

$$\int_0^\infty e^{-\lambda t} \pi(d\lambda) = \frac{1}{\Gamma(2\overline{H})} \int_0^\infty e^{-xt} e^{-x} x^{2\overline{H}-1} dx = (t+1)^{-2\overline{H}}.$$

Note in particular that $\tilde{V}$ exhibits second order long range dependence if $H \in (\frac{1}{2}, 1)$ for $H := 1 - \overline{H}$.

5.3.2 Volatility modulated supOU processes

In a next step, we extend the class of supOU processes and introduce a new concept which allow us to introduce long memory and stochastic volatility simultaneously. As before, we present the general case where we have both a stochastic proportional and a stochastic time change.

Note here that the stochastic proportional can be easily included in the supOU framework, whereas the time change requires some additional work. More precisely, the time change approach leads to integration with respect to a random measure more general than a Lévy basis.

Recall that the stochastic volatility process is defined as

$$V_t = \int_{-\infty}^t e^{-\lambda(t-s)} \upsilon_{\lambda s-} dL_{\tau_{\lambda s}}.$$

Let T be the random measure associated with the stochastic process τ, so that for intervals $(a, b]$ we have $T((a, b]) = \tau(b) - \tau(a)$. (If τ was a Lévy subordinator, then T would be the corresponding Lévy basis.) We introduce a random measure M on $\mathbb{R} \times \mathbb{R}_+$ characterised by requiring that M conditionally on T is the Lévy basis on $\mathbb{R} \times \mathbb{R}_+$ having characteristic quadruplet

$$(0, 0, \nu(dx), T(\lambda dt)\pi(d\lambda)). \tag{19}$$

Here ν and π are as above and the control measure is $T(\lambda dt)\pi(d\lambda)$. (This construction is analogous that of extended subordination by meta–time changes of Lévy bases as introduced in Barndorff-Nielsen (2010) and Barndorff-Nielsen & Pedersen (2011).)

Then, define the process $\tilde{V}$ by

$$\tilde{V}_t = \int_0^\infty \int_{-\infty}^t e^{-\lambda(t-s)} \upsilon_{\lambda s-} M(ds, d\lambda).$$

That this determines a well–defined strictly stationary process can be verified by calculating the characteristic functional of $\tilde{V}$.

Under square integrability, the conditional autocovariance function given υ and τ takes the form

$$\mathbb{E}\left((\tilde{V}_0 - E(\tilde{V}_0|\upsilon,\tau))(\tilde{V}_t - E(\tilde{V}_t|\upsilon,\tau)) \Big| \upsilon, \tau \right) = \kappa_2 \int_0^\infty \int_{-\infty}^0 e^{-\lambda t + 2\lambda s} \upsilon^2_{\lambda s-} T(\lambda ds)\pi(d\lambda)$$

$$= \kappa_2 \int_0^\infty e^{-\lambda t}\pi(d\lambda) \int_{-\infty}^0 e^{2\lambda s}\upsilon_{\lambda s-}^2 T(\lambda ds) = \int_0^\infty e^{-\lambda t}\pi(d\lambda) \underbrace{\kappa_2 \int_{-\infty}^0 e^{2u}\upsilon_{u-}^2 T(du)}_{:=X^{(2)}}$$

$$= X^{(2)} \int_0^\infty e^{-\lambda t}\pi(d\lambda) = X^{(2)}\widehat{\pi}(\lambda),$$

where the random variable $X^{(2)}$ and its law do not depend on λ. As before, $\widehat{\pi}(\lambda)$ denotes the Laplace transform of π.

As in the example above, long memory of $\tilde{V}$ can be obtained by choosing π to be the probability measure of the Gamma law.

Hence we have seen that the volatility modulated supOU process $\widetilde{V}$ can account for long memory and stochastic volatility simultaneously.

Note here that the novel contribution is the way how long memory and the time change are combined: We have seen that additional volatility and long memory can be introduced in a non–Gaussian OU process by an extended subordination approach, where the additional volatility enters through a time change τ, with associated random measure T, and the long memory can be obtained by a suitable choice of the probability measure π. These two measures have to be combined, as described in the characteristic quadruplet defined in (19), in order to obtain both long memory and additional stochastic volatility.

6 Concluding remarks

This paper has introduced a new class of stochastic volatility models which is given by volatility modulated non–Gaussian Ornstein–Uhlenbeck (VMOU) processes. We have shown that the new model class is highly analytically tractable and, in particular, we have derived an explicit formula for the integrated squared volatility process, which plays a key role in determining an explicit formula for the variance risk premium.

Next, we have shown that the additional volatility of volatility component can be used to introduce the leverage effect in a new way. Also, we have developed a new methodology for allowing for long memory *and* volatility of volatility simultaneously: This can be done by combining the concepts of extended subordination of Lévy bases (or of more general random measures) and of randomisation of the memory parameter of the OU process through a suitable choice of the characteristic quadruplet.

Another key result we have established in this paper is the fact that the stochastic volatility of volatility component solely determines the stochastic dynamics of the variance risk premium if the change of measure is structure–preserving. Given the empirical evidence that the variance risk premium has stochastic dynamics, including a stochastic volatility of volatility component into a stochastic volatility model is hence a modelling choice which leads to rather explicit dynamics of the variance risk premium. Clearly, there are various natural extensions of our new model. For instance, we will address multivariate extensions of VMOU processes and, also, superpositions of such multivariate processes in future research. Multivariate OU processes and their superpositions have recently been introduced by Barndorff-Nielsen & Stelzer (2011) and have been applied as multivariate stochastic volatility models by Barndorff-Nielsen & Stelzer (2012), Muhle-Karbe et al. (2012), Pigorsch & Stelzer (2009). Furthermore, we plan to address multivariate extensions of the VMOU. In particular, for the time change case, future research will be based on the new concept of multivariate subordination introduced by Barndorff-Nielsen (2010) and Barndorff-Nielsen & Pedersen (2011).

Acknowledgement

We would like to thank the editor Eric Renault and two anonymous referees for very valuable and constructive comments on an earlier draft of this paper.

This work was supported by the Center for Research in Econometric Analysis of Time Series, CREATES, funded by the Danish National Research Foundation.

A Proofs

A.1 Second order structure

Here we study various conditional and unconditional moments of V and we derive its autocorrelation function. In the following, we will often present the results only and omit the rather lengthy proofs since they consist of straightforward computations.

In order to simplify the exposition, we fix the following notation. For $i \in \mathbb{N}$, we denote the ith cumulant of the Lévy subordinator L_t by $\kappa_i(L_t)$ (provided it exists). Clearly, we have $\kappa_i(L_t) = t\kappa_i(L_1)$ and, in particular, we write $\kappa_i := \frac{1}{\lambda}\kappa_i(L_\lambda) = \kappa_i(L_1)$.

We will carry out all computations for a general stationary volatility process υ and an absolutely continuous time change τ with stationary density process ξ. In the following, we will write $\gamma(h) = Cov(\upsilon_{t-}\xi_{t-}, \upsilon_{(t+h)-}\xi_{(t+h)-})$ for $h > 0$. Recall also that we assume mutual independence of L, υ and τ.

First, we compute the moments of V, when we condition on υ and τ. Clearly, if υ and τ were deterministic, these results would also hold unconditionally.

Throughout this section, we will use the following notation for the conditional expectation: $\mathbb{E}^{\upsilon,\tau}(\cdot) := \mathbb{E}(\cdot|\upsilon,\tau)$, $Var^{\upsilon,\tau}(\cdot) := Var(\cdot|\upsilon,\tau)$, $Cov^{\upsilon,\tau}(\cdot,\cdot) := Cov(\cdot,\cdot|\upsilon,\tau)$ and $Cor^{\upsilon,\tau}(\cdot,\cdot) := Cor(\cdot,\cdot|\upsilon,\tau)$.

Proposition 17 *The conditional mean, variance and covariance are given by*

$$\mathbb{E}^{\upsilon,\tau}(V_t) = e^{-\lambda t}\mathbb{E}^{\upsilon,\tau}(V_0) + \lambda\kappa_1(L_1)\int_0^t e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s}ds,$$

$$Var^{\upsilon,\tau}(V_t) = e^{-2\lambda t}Var^{\upsilon,\tau}(V_0) + \lambda\kappa_2(L_1)\int_0^t e^{-2\lambda(t-s)}\upsilon_{\lambda s}^2\xi_{\lambda s}ds,$$

$$Cov^{\upsilon,\tau}(V_t, V_{t+h}) = e^{-\lambda h}Var^{\upsilon,\tau}(V_t).$$

Proof of Proposition 17 For the mean, we have

$$\mathbb{E}^{\upsilon,\tau}(V_t) = e^{-\lambda t}\mathbb{E}^{\upsilon,\tau}(V_0) + \lambda\kappa_1(L_1)\int_0^t e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s}ds.$$

For the second moment, we get from Itô's formula

$$V_t^2 - V_0^2 = -2\lambda\int_0^t V_s^2 ds + 2\int_0^t V_s\upsilon_{\lambda s-}dL_{\tau_{\lambda s}} + \int_0^t \upsilon_{\lambda s-}^2 d[L]_{\tau_{\lambda s}}.$$

Taking the conditional expectation, we get

$$\begin{aligned}&\mathbb{E}^{\upsilon,\tau}(V_t^2) - \mathbb{E}^{\upsilon,\tau}(V_0^2)\\ &= -2\lambda\int_0^t \mathbb{E}^{\upsilon,\tau}(V_s^2)ds + 2\lambda\kappa_1(L_1)\int_0^t \mathbb{E}^{\upsilon,\tau}(V_s\upsilon_{\lambda s}\xi_{\lambda s})ds + \lambda\kappa_2(L_1)\int_0^t \mathbb{E}^{\upsilon,\tau}(\upsilon_{\lambda s}^2\xi_{\lambda s})ds\end{aligned}$$

$$= -2\lambda \int_0^t \mathbb{E}^{\upsilon,\tau}(V_s^2)ds + 2\lambda\kappa_1(L_1)\int_0^t \mathbb{E}^{\upsilon,\tau}(V_s)\upsilon_{\lambda s}\xi_{\lambda s}ds + \lambda\kappa_2(L_1)\int_0^t \upsilon_{\lambda s}^2\xi_{\lambda s}ds.$$

Hence,

$$d\mathbb{E}^{\upsilon,\tau}(V_t^2) = -2\lambda\mathbb{E}^{\upsilon,\tau}(V_t^2)dt + \left(2\lambda\kappa_1(L_1)\mathbb{E}^{\upsilon,\tau}(V_t)\upsilon_{\lambda t}\xi_{\lambda s} + \lambda\kappa_2(L_1)\upsilon_{\lambda t}^2\xi_{\lambda s}\right)dt.$$

We solve the linear stochastic differential equation and obtain

$$\mathbb{E}^{\upsilon,\tau}(V_t^2) = e^{-2\lambda t}\mathbb{E}^{\upsilon,\tau}(V_0^2) + \int_0^t e^{-2\lambda(t-s)}\Xi_s ds,$$

where

$$\begin{aligned}\Xi_s &:= 2\lambda\kappa_1(L_1)\mathbb{E}^{\upsilon,\tau}(V_s)\upsilon_{\lambda s}\xi_{\lambda s} + \lambda\kappa_2(L_1)\upsilon_{\lambda s}^2\xi_{\lambda s}\\ &= 2\lambda\kappa_1(L_1)e^{-\lambda s}\mathbb{E}^{\upsilon,\tau}(V_0)\upsilon_{\lambda s}\xi_{\lambda s} + 2\lambda^2\kappa_1^2(L_1)\upsilon_{\lambda s}\xi_{\lambda s}\int_0^s e^{-\lambda(s-u)}\upsilon_{\lambda u}\xi_{\lambda u}du + \lambda\kappa_2(L_1)\upsilon_{\lambda s}^2\xi_{\lambda s}.\end{aligned}$$

Consequently, we have

$$\begin{aligned}\mathbb{E}^{\upsilon,\tau}(V_t^2) &= e^{-2\lambda t}\mathbb{E}^{\upsilon,\tau}(V_0^2) + \int_0^t e^{-2\lambda(t-s)}\Xi_s ds\\ &= e^{-2\lambda t}\mathbb{E}^{\upsilon,\tau}(V_0^2) + 2\lambda\kappa_1(L_1)\mathbb{E}^{\upsilon,\tau}(V_0)\int_0^t \underbrace{e^{-2\lambda(t-s)}e^{-\lambda s}}_{=e^{-\lambda t}e^{-\lambda(t-s)}}\upsilon_{\lambda s}\xi_{\lambda s}ds\\ &\quad + 2\lambda^2\kappa_1^2(L_1)\underbrace{\int_0^t e^{-2\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s}\int_0^s e^{-\lambda(s-u)}\upsilon_{\lambda u}\xi_{\lambda u}duds}_{=e^{-2\lambda t}\int_0^t e^{\lambda s}\upsilon_{\lambda s}\xi_{\lambda s}\int_0^s e^{\lambda u}\upsilon_{\lambda u}\xi_{\lambda u}duds}\\ &\quad + \lambda\kappa_2(L_1)\int_0^t e^{-2\lambda(t-s)}\upsilon_{\lambda s}^2\xi_{\lambda s}ds.\end{aligned}$$

Hence,

$$Var^{\upsilon,\tau}(V_t) = e^{-2\lambda t}Var^{\upsilon,\tau}(V_0) + \lambda\kappa_2(L_1)\int_0^t e^{-2\lambda(t-s)}\upsilon_{\lambda s}^2\xi_{\lambda s}ds.$$

Similar computations lead to the result for the covariance. In particular, using the representation results for the increments of V, we get

$$\begin{aligned}\mathbb{E}^{\upsilon,\tau}(V_tV_{t+h}) &= e^{-\lambda h}\mathbb{E}^{\upsilon,\tau}(V_t^2) + \lambda\kappa_1(L_1)\mathbb{E}^{\upsilon,\tau}(V_t)\int_t^{t+h} e^{-\lambda(t+h-s)}\upsilon_{\lambda s}\xi_{\lambda s}ds\\ &= e^{-\lambda h}\left(\mathbb{E}^{\upsilon,\tau}(V_t^2) + \lambda\kappa_1(L_1)\mathbb{E}^{\upsilon,\tau}(V_t)\int_t^{t+h} e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s}ds\right)\\ &= e^{-\lambda h}\left(\mathbb{E}^{\upsilon,\tau}(V_t^2) + \lambda\kappa_1(L_1)e^{-\lambda t}\mathbb{E}^{\upsilon,\tau}(V_0)\int_t^{t+h} e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s}ds\right.\\ &\qquad \left. + \lambda^2\kappa_1^2(L_1)\int_0^t e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s}ds\int_t^{t+h} e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s}ds\right).\end{aligned}$$

□

Next, we compute the unconditional mean, variance and covariance of the VMOU process V.

Proposition 18 *Let* $V_t = V_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} \upsilon_{\lambda s-} dL_{\tau_{\lambda s}}$. *Then, for* $h > 0$,

$$
\begin{aligned}
\mathbb{E}(V_t) &= e^{-\lambda t}\mathbb{E}(V_0) + \kappa_1(L_1)\mathbb{E}(\upsilon_0)\mathbb{E}(\xi_0)(1-e^{-\lambda t}),\\
Var(V_t) &= e^{-2\lambda t} Var(V_0) + 2e^{-\lambda t}\kappa_1(L_1) Cov\left(\mathbb{E}^{\upsilon,\tau}(V_0), \lambda \int_0^t e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s} ds\right)\\
&\quad + \kappa_1^2 \int_0^t \int_0^t e^{-\lambda(t-s)} e^{-\lambda(t-u)} \lambda^2 \gamma(|\lambda s - \lambda u|) ds du\\
&\quad + \kappa_2(L_1)\mathbb{E}(\upsilon_0^2)\mathbb{E}(\xi_0)\frac{1}{2}(1-e^{-2\lambda t}),\\
Cov(V_t, V_{t+h}) &= e^{-\lambda h}\left\{ Var(V_t) + \kappa_1(L_1) e^{-\lambda t} Cov\left[\mathbb{E}^{\upsilon,\tau}(V_0), \int_t^{t+h} \lambda e^{-\lambda(t-s)} \upsilon_{\lambda s}\xi_{\lambda s} ds\right]\right.\\
&\quad \left. + \kappa_1^2(L_1) \int_0^t \int_t^{t+h} \lambda e^{-\lambda(t-s)} \lambda e^{-\lambda(t-u)} \gamma(|\lambda s - \lambda u|) du ds \right\}.
\end{aligned}
$$

Proof The results follow essentially from Proposition 17. We only focus on the computations of the variance here. Recall that

$$Var(V_t) = \mathbb{E}(Var^{\upsilon,\tau}(V_t)) + Var(\mathbb{E}^{\upsilon,\tau}(V_t)).$$

The first term is straightforward to compute. For the second term we get

$$Var(\mathbb{E}^{\upsilon,\tau}(V_t)) = \mathbb{E}((\mathbb{E}^{\upsilon,\tau}(V_t))^2) - (\mathbb{E}(\mathbb{E}^{\upsilon,\tau}(V_t)))^2.$$

Then

$$
\begin{aligned}
\mathbb{E}((\mathbb{E}^{\upsilon,\tau}(V_t))^2) &= \mathbb{E}\left[e^{-2\lambda t}\left(\mathbb{E}^{\upsilon,\tau}(V_0)\right)^2 + 2e^{-\lambda t}\lambda\kappa_1(L_1)\mathbb{E}^{\upsilon,\tau}(V_0)\int_0^t e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s} ds\right.\\
&\quad \left. + \lambda^2\kappa_1^2(L_1)\left(\int_0^t e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s} ds\right)^2\right]\\
&= e^{-2\lambda t}\mathbb{E}\left[\left(\mathbb{E}^{\upsilon,\tau}(V_0)\right)^2\right] + 2e^{-\lambda t}\lambda\kappa_1(L_1)\mathbb{E}\left[\mathbb{E}^{\upsilon,\tau}(V_0)\int_0^t e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s} ds\right]\\
&\quad + \lambda^2\kappa_1^2(L_1)\mathbb{E}\left[\left(\int_0^t e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s} ds\right)^2\right].
\end{aligned}
$$

Then, we get that

$$
\begin{aligned}
Var(V_t) &= e^{-2\lambda t} Var(V_0) + 2e^{-\lambda t}\kappa_1(L_1) Cov\left(\mathbb{E}^{\upsilon,\tau}(V_0), \lambda\int_0^t e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s} ds\right)\\
&\quad + \kappa_1^2 Var\left(\int_0^t \lambda e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s} ds\right) + \kappa_2(L_1)\mathbb{E}(\upsilon_0^2)\mathbb{E}(\xi_0)\underbrace{\int_0^t \lambda e^{-2\lambda(t-s)} ds}_{=\frac{1}{2}(1-e^{-2\lambda t})},
\end{aligned}
$$

where

$$Var\left(\int_0^t \lambda e^{-\lambda(t-s)}\upsilon_{\lambda s}\xi_{\lambda s} ds\right) = e^{-2\lambda t}\int_0^t\int_0^t \lambda^2 e^{\lambda s} e^{\lambda u}\gamma(|\lambda s - \lambda u|) ds du$$

$$= \int_0^{\lambda t} \int_0^{\lambda t} e^v e^w \gamma(|v-w|) dv dw.$$

For the covariance, we get

$$\begin{aligned}
Cov(V_t, V_{t+h}) &= \mathbb{E}(\mathbb{E}^{\upsilon,\tau}(V_t V_{t+h})) - \mathbb{E}(\mathbb{E}^{\upsilon,\tau}(V_t))\mathbb{E}(\mathbb{E}^{\upsilon,\tau}(V_{t+h})) \\
&= e^{-\lambda h} \left\{ \mathbb{E}[\mathbb{E}^{\upsilon,\tau}(V_t^2)] + \lambda \kappa_1(L_1) e^{-\lambda t} \mathbb{E}\left[\mathbb{E}^{\upsilon,\tau}(V_0) \int_t^{t+h} e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds \right] \right. \\
&\quad \left. + \lambda^2 \kappa_1^2(L_1) \mathbb{E}\left[\int_0^t e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds \int_t^{t+h} e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds \right] \right\} \\
&\quad - e^{-\lambda h} \left\{ (\mathbb{E}(V_t))^2 + E(V_t) \kappa_1(L_1) \mathbb{E}(\upsilon_0) \mathbb{E}(\xi_0) \int_t^{t+h} e^{-\lambda(t-s)} ds \right\} \\
&= e^{-\lambda h} \left\{ Var(V_t) + (\mathbb{E}(V_t))^2 + \lambda \kappa_1(L_1) e^{-\lambda t} \mathbb{E}\left[\mathbb{E}^{\upsilon,\tau}(V_0) \int_t^{t+h} e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds \right] \right. \\
&\quad + \lambda^2 \kappa_1^2(L_1) \mathbb{E}\left[\int_0^t e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds \int_t^{t+h} e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds \right] \Bigg) \\
&\quad \left. - (\mathbb{E}(V_t))^2 - \mathbb{E}(V_t) \kappa_1(L_1) \mathbb{E}(\upsilon_0) \mathbb{E}(\xi_0) \int_t^{t+h} \lambda^{-\lambda(t-s)} ds \right\} \\
&= e^{-\lambda h} \left\{ Var(V_t) + \kappa_1(L_1) e^{-\lambda t} Cov\left[\mathbb{E}^{\upsilon,\tau}(V_0), \int_t^{t+h} \lambda e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds \right] \right. \\
&\quad \left. + \kappa_1^2(L_1) Cov\left[\int_0^t \lambda e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds, \int_t^{t+h} \lambda e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds \right] \right\},
\end{aligned}$$

where

$$\begin{aligned}
& Cov\left[\int_0^t \lambda e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds, \int_t^{t+h} \lambda e^{-\lambda(t-s)} \upsilon_{\lambda s} \xi_{\lambda s} ds \right] \\
&= e^{-2\lambda t} \int_0^t \int_t^{t+h} \lambda^2 e^{\lambda s} e^{\lambda u} \gamma(|\lambda s - \lambda u|) ds du \\
&= \int_0^{\lambda t} \int_{\lambda t}^{\lambda(t+h)} e^v e^w \gamma(|v-w|) dv dw.
\end{aligned}$$

□

Corollary 19 *The mean, variance and autocovariance of the stationary process V are given by*

$$\begin{aligned}
\mathbb{E}(V_t) &= \kappa_1(L_1) \mathbb{E}(\upsilon_0) \mathbb{E}(\xi_0), \\
Var(V_t) &= \frac{1}{2} \kappa_2(L_1) \mathbb{E}\left(\upsilon_0^2\right) \mathbb{E}(\xi_0) + \kappa_1^2(L_1) \int_{-\infty}^0 \int_{-\infty}^0 e^x e^y \gamma(|x-y|) dx dy, \\
Cov(V_t, V_{t+h}) &= e^{-\lambda h} \left(\frac{1}{2} \kappa_2(L_1) \mathbb{E}\left(\upsilon_0^2\right) \mathbb{E}(\xi_0) + \kappa_1^2(L_1) \int_{-\infty}^0 \int_{-\infty}^{\lambda h} e^x e^y \gamma(|x-y|) dx dy \right).
\end{aligned}$$

Proof The results follow directly from the previous results, where we used that the initial value for the stationary process is given by $V_0 = \int_{-\infty}^0 e^{\lambda s} \upsilon_{\lambda s-} dL_{\tau_{\lambda s}}$. □

A.2 The leverage effect

Proof of the results in Section 2.6 Note that $P_0 = 0$ and $\mathbb{E}(P_t) = 0$. From Itô's product rule, we get that

$$\begin{aligned} P_t V_t &= \int_0^t V_{s-} dP_s + \int_0^t P_{s-} dV_s + [P,V]_t \\ &= \int_0^t V_{s-} dP_s - \lambda \int_0^t P_s V_s ds + \int_0^t P_{s-} \upsilon_{\lambda s-} dL_{\lambda s} + [P,V]_t. \end{aligned}$$

Taking expectations, we get

$$Cov(P_t, V_t) = \mathbb{E}(P_t V_t) = -\lambda \int_0^t \mathbb{E}(P_s V_s) ds + \lambda \mathbb{E}(L_1) \int_0^t \mathbb{E}\left(P_s \upsilon_{\lambda s}\right) ds.$$

The above equation is an integral equation, which can be solved as soon as we have computed the second term on the right hand side. We do that by applying Itô's product rule again and obtain

$$\begin{aligned} \mathbb{E}\left(P_u \upsilon_{\lambda u}\right) &= \mathbb{E}\left(\int_0^u P_s d\upsilon_{\lambda s}\right) + \mathbb{E}\left(\int_0^u \sqrt{V_{s-}} d[W, \upsilon_\lambda]_s\right) \\ &= -\alpha\lambda \int_0^u \mathbb{E}(P_s \upsilon_{\lambda s}) ds + \widetilde{\rho}\sqrt{\lambda}\gamma \int_0^u \mathbb{E}\left(\sqrt{V_s}\sqrt{\upsilon_{\lambda s}}\right) ds, \end{aligned}$$

which is yet another integral equation. Since V and υ_λ are strictly positive (provided the Feller condition $2\alpha\beta > \gamma^2$ and $\upsilon_0 > 0$ holds) , we have that $g(s) := \mathbb{E}\left(\sqrt{V_s^{(1)}}\sqrt{\upsilon_{\lambda s}}\right)$, for a strictly positive function g. Solving the differential equation

$$\frac{d}{du}\mathbb{E}\left(P_u \upsilon_{\lambda u}\right) = -\alpha\lambda \mathbb{E}(P_u \upsilon_{\lambda u}) + \widetilde{\rho}\sqrt{\lambda}\gamma g(u),$$

we get

$$\mathbb{E}\left(P_u \upsilon_{\lambda u}\right) = \exp(-\alpha\lambda u)\left(\int_0^u \widetilde{\rho}\sqrt{\lambda}\gamma g(s)\exp(\alpha\lambda s) ds\right) =: \widetilde{\rho}\widetilde{g}(u) \not\equiv 0,$$

for a strictly positive function $\widetilde{g}$. Next, we define $G(u) := \lambda \mathbb{E}(L_1)\mathbb{E}\left(P_u \upsilon_{\lambda u}\right)$. Solving

$$\frac{d}{dt}\mathbb{E}(P_t V_t) = -\lambda \mathbb{E}(P_t V_t) + G(t),$$

we obtain

$$\mathbb{E}(P_t V_t) = \exp(-\lambda t)\int_0^t G(x) e^{\lambda x} dx \not\equiv 0.$$

Further, Itô's formula leads to the following result for higher moments of order $n \in \mathbb{R}, n \geq 2$ (provided they exist):

$$\mathbb{E}\left(P_t^n\right) = \frac{n(n-1)}{2}\int_0^t \mathbb{E}\left(P_s^{n-2} V_s\right) ds.$$

In particular, we have

$$Cov(P_t, P_t^2) = \mathbb{E}\left(P_t^3\right) = 3\int_0^t \mathbb{E}(P_s V_s) ds = 3\lambda \mathbb{E}(L_1)\int_0^t e^{-\lambda s}\int_0^s e^{\lambda u}\mathbb{E}\left(P_u \upsilon_{\lambda u}\right) du ds \not\equiv 0.$$

□

A.3 The Laplace transform of integrated squared volatility

Proof of Proposition 5 Since the integrated variance appears in the asset price formula, we compute the Laplace transform of the integrated variance first. In particular, we focus on the *conditional* integrated variance over the interval $[t, t+h]$ given $\mathcal{F}_t$ for $t, h \geq 0$. We use the notation $\mathbb{E}_t^{\mathbb{Q}}(\cdot) := \mathbb{E}^{\mathbb{Q}}(\cdot|\mathcal{F}_t)$. Further, using (6), we get

$$\begin{aligned}
\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\int_t^{t+h} V_s ds\right)\right] &= \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\left(\epsilon_\lambda(h)V_t + \int_t^{t+h}\epsilon_\lambda(t+h-s)dL_{\lambda s}^{\upsilon,\tau}\right)\right)\right] \\
&= \exp\left(\theta\epsilon_\lambda(h)V_t\right)\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)dL_{\lambda s}^{\upsilon,\tau}\right)\right)\right].
\end{aligned}$$

We define the following σ–algebras $\mathcal{G}^{(1)} := \sigma(\upsilon_{\lambda s} : s \leq t+h) \cup \mathcal{F}_t$ and $\mathcal{G}^{(2)} := \sigma(\tau_{\lambda s} : s \leq t+h) \cup \mathcal{F}_t$. Then the Laplace transformation is given by

$$\begin{aligned}
&\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)dL_{\lambda s}^{\upsilon,\tau}\right)\right)\right] = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)\upsilon_{\lambda s-}dL_{\tau_{\lambda s}}\right)\right)\right] \\
&= \mathbb{E}_t^{\mathbb{Q}}\left\{\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)\upsilon_{\lambda s-}dL_{\tau_{\lambda s}}\right)\right)\middle|\mathcal{G}^{(1)},\mathcal{G}^{(2)}\right]\right\} \\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\int_t^{t+h}\chi_L(\theta\epsilon_\lambda(t+h-s)\upsilon_{\lambda s-})\tau(\lambda ds)\right)\right] \\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\int_t^{t+h}\chi_L(\theta\epsilon_\lambda(t+h-s)\upsilon_{\lambda s-})\xi_{\lambda s}\lambda ds\right)\right],
\end{aligned}$$

where χ_L denotes the log–transformed Laplace transform of L.

Let $\mathcal{G}^{(3)} := \sigma(L_{\lambda s}^{\upsilon,\tau}|s \leq t+h) \cup \mathcal{F}_t$. For the conditional distribution of Y, we get

$$\begin{aligned}
\phi(\theta) &:= \mathbb{E}_t^{\mathbb{Q}}(\exp(\theta Y_{t+h})) = \exp(\theta Y_t)\mathbb{E}_t^{\mathbb{Q}}(\exp(\theta(Y_{t+h}-Y_t))) \\
&= \exp(\theta Y_t)\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\left(\mu h + \beta\int_t^{t+h}V_s ds + \int_t^{t+h}\sqrt{V_{s-}}dW_s\right)\right)\right] \\
&= \exp(\theta Y_t)\mathbb{E}_t^{\mathbb{Q}}\left[\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\left(\mu h + \beta\int_t^{t+h}V_s ds + \int_t^{t+h}\sqrt{V_{s-}}dW_s\right)\right)\middle|\mathcal{G}^{(1)},\mathcal{G}^{(2)},\mathcal{G}^{(3)}\right]\right] \\
&= \exp(\theta Y_t)\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\left(\mu h + \beta\int_t^{t+h}V_s ds\right)\right)\right. \\
&\qquad \left.\underbrace{\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\int_t^{t+h}\sqrt{V_s}dW_s\right)\middle|\mathcal{G}^{(1)},\mathcal{G}^{(2)},\mathcal{G}^{(3)}\right]}_{=\exp\left(\frac{\theta^2}{2}\int_t^{t+h}V_s ds\right)}\right] \\
&= \exp(\theta Y_t)\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\theta\mu h + \left(\theta\beta + \frac{\theta^2}{2}\right)\int_t^{t+h}V_s ds\right)\right] \\
&= \exp\left(\theta Y_t + \theta\mu h + \left(\beta\theta + \frac{\theta^2}{2}\right)\epsilon_\lambda(h)V_t\right)\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\int_t^{t+h}f(s,\theta)dL_{\tau_{\lambda s}}\right)\right] \\
&= \exp\left(\theta Y_t + \theta\mu h + \left(\beta\theta + \frac{\theta^2}{2}\right)\epsilon_\lambda(h)V_t\right)\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\lambda\int_t^{t+h}\chi_L(f(s,\theta))d\tau_{\lambda s}\right)\right]
\end{aligned}$$

$$= \exp\left(\theta Y_t + \theta\mu h + \left(\beta\theta + \frac{\theta^2}{2}\right)\epsilon_\lambda(h)V_t\right)\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\lambda\int_t^{t+h}\chi_L(f(s,\theta))\xi_{\lambda s}\lambda ds\right)\right],$$

where $f(s,\theta) := (\beta\theta + \theta^2/2)\epsilon_\lambda(t+h-s)\upsilon_{\lambda s-}$. □

A.4 Change of measure and variance risk premium

Proof of Proposition 6 We apply the following Bayes rule, see Karatzas & Shreve (1991, p.193). For any $\mathcal{F}_t$ measurable random variable ζ with $\mathbb{E}^{\mathbb{Q}_{t+h}}|\zeta| < \infty$ and for $0 \le s \le t \le t+h$, we have

$$\mathbb{E}_s^{\mathbb{Q}_{t+h}}(\zeta) = \frac{1}{Z_s}\mathbb{E}_s(\zeta Z_t),$$

and, hence

$$\mathbb{E}_s(\zeta) - \mathbb{E}_s^{\mathbb{Q}_{t+h}}(\zeta) = \mathbb{E}_s\left(\zeta\left(1 - \frac{Z_t}{Z_s}\right)\right).$$

As a particular case, we get for the approximated variance risk premium, since $[Y]_{[t,t+h]}$ is $\mathcal{F}_{t+h}$ measurable,

$$VRP_{t,t+h}^{\mathbb{Q}} = \mathbb{E}_t\left([Y]_{[t,t+h]}\left(1 - \frac{Z_{t+h}}{Z_t}\right)\right).$$

Recall that

$$[Y]_{[t,t+h]} = \int_t^{t+h}\sigma_s^2 ds + \sum_{t\le s\le t+h}(\Delta J_s)^2 = [Y]^c_{[t,t+h]} + [Y]^d_{[t,t+h]},$$

where $[Y]^c_{[t,t+h]}$ denotes the continuous part of the quadratic variation and $[Y]^d_{[t,t+h]}$ the jump part. Next we plug in the explicit formula for V^+, see (6), and we obtain

$$[Y]^c_{[t,t+h]} = V^+_{t+h} - V^+_t = \epsilon_\lambda(h)V_t + \int_t^{t+h}\epsilon_\lambda(t+h-s)dL^{\upsilon,\tau}_{\lambda s}.$$

Hence

$$\begin{aligned}\mathbb{E}_t\left([Y]^c_{[t,t+h]}\left(1 - \frac{Z_{t+h}}{Z_t}\right)\right) &= \epsilon_\lambda(h)V_t\underbrace{\mathbb{E}_t\left(\left(1 - \frac{Z_{t+h}}{Z_t}\right)\right)}_{=0} \\ &\quad + \mathbb{E}_t\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)dL^{\upsilon,\tau}_s\left(1 - \frac{Z_{t+h}}{Z_t}\right)\right) \\ &= \mathbb{E}_t\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)dL^{\upsilon,\tau}_s\left(1 - \frac{Z_{t+h}}{Z_t}\right)\right).\end{aligned}$$

For the jump part of the quadratic variation, we get

$$\mathbb{E}_t\left([Y]^d_{[t,t+h]}\left(1 - \frac{Z_{t+h}}{Z_t}\right)\right) = \mathbb{E}_t\left(\sum_{t\le s\le t+h}(\Delta J_s)^2\left(1 - \frac{Z_{t+h}}{Z_t}\right)\right).$$

□

Proof of Proposition 8 Since the jumps come from a Lévy process, the conditional expectation $\mathbb{E}_t$ equals the unconditional one and we get

$$VRP^{\mathbb{Q},d}_{t,t+h} = \frac{1}{h}\left[\mathbb{E}_t\left(\sum_{t\le s\le t+h}(\Delta J_s)^2\right) - \mathbb{E}^{\mathbb{Q}}_t\left(\sum_{t\le s\le t+h}(\Delta J_s)^2\right)\right]$$
$$= \int_{\mathbb{R}} x^2\nu_J(dx) - \int_{\mathbb{R}} x^2\nu^{\mathbb{Q}}_J(dx),$$

where $\nu_J(\cdot)$ and $\nu^{\mathbb{Q}}_J(\cdot)$ are the Lévy measures of the Lévy process J under $\mathbb{P}$ and under $\mathbb{Q}$, respectively. □

Proof of Proposition 9 Due to the independence of υ and τ, we get

$$\mathbb{E}_t\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)dL^{\upsilon,\tau}_{\lambda s}\right) = \lambda\mathbb{E}(L_1)\int_t^{t+h}\epsilon_\lambda(t+h-s)\mathbb{E}_t\left(\upsilon_{\lambda s-}\right)\mathbb{E}_t\left(\xi_{\lambda s-}\right)ds.$$

Under a structure–preserving measure change, we have

$$\mathbb{E}^{\mathbb{Q}}_t\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)dL^{\upsilon,\tau}_{\lambda s}\right) = \lambda\mathbb{E}^{\mathbb{Q}}(L_1)\int_t^{t+h}\epsilon_\lambda(t+h-s)\mathbb{E}^{\mathbb{Q}}_t\left(\upsilon_{\lambda s-}\right)\mathbb{E}^{\mathbb{Q}}_t\left(\xi_{\lambda s-}\right)ds.$$

Hence

$$\mathbb{E}_t\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)dL^{\upsilon,\tau}_{s}\right) - \mathbb{E}^{\mathbb{Q}}_t\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)dL^{\upsilon,\tau}_{s}\right)$$
$$= \lambda\int_t^{t+h}\epsilon_\lambda(t+h-s)\Big(\mathbb{E}(L_1)\mathbb{E}_t\left(\upsilon_{\lambda s-}\right)\mathbb{E}_t\left(\xi_{\lambda s-}\right) - \mathbb{E}^{\mathbb{Q}}(L_1)\mathbb{E}^{\mathbb{Q}}_t\left(\upsilon_{\lambda s-}\right)\mathbb{E}^{\mathbb{Q}}_t\left(\xi_{\lambda s-}\right)\Big)ds$$
$$= \lambda\int_t^{t+h}\epsilon_\lambda(t+h-s)\Big(\mathbb{E}(L_1)\mathbb{E}_t\left(\upsilon_{\lambda s-}\xi_{\lambda s-}\right) - \mathbb{E}^{\mathbb{Q}}(L_1)\mathbb{E}^{\mathbb{Q}}_t\left(\upsilon_{\lambda s-}\xi_{\lambda s-}\right)\Big)ds.$$

Hence, we see clearly that the stochastic proportional υ and the density of the time change ξ play the same role in determining the dynamics of the variance risk premium. Now, we define $\eta_t = \upsilon_{\lambda t-}\xi_{\lambda t-}$. Then

$$\mathbb{E}_t\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)dL^{\upsilon,\tau}_{s}\right) - \mathbb{E}^{\mathbb{Q}}_t\left(\int_t^{t+h}\epsilon_\lambda(t+h-s)dL^{\upsilon,\tau}_{s}\right)$$
$$= \eta^{\upsilon,\tau}_t\Big(\mathbb{E}(L_1) - \mathbb{E}^{\mathbb{Q}}(L_1)\Big)(h-\epsilon_\lambda(h))$$
$$+ \lambda\int_t^{t+h}\epsilon_\lambda(t+h-s)\Big(\mathbb{E}(L_1)\mathbb{E}_t\left(\eta^{\upsilon,\tau}_s - \eta^{\upsilon,\tau}_t\right) - \mathbb{E}^{\mathbb{Q}}(L_1)\mathbb{E}^{\mathbb{Q}}_t\left(\eta^{\upsilon,\tau}_s - \eta^{\upsilon,\tau}_t\right)\Big)ds.$$

Note that L is a subordinator and, hence, $\mathbb{E}(L_1) > 0$. Also, under the structure–preserving change of measure, the predictable compensator of L changes and, hence, $\mathbb{E}(L_1) - \mathbb{E}^{\mathbb{Q}}(L_1) \neq 0$. This concludes the proof. □

Proof of Corollary 10 Throughout the proof we assume that $s \geq t$. Further, we skip the superscript and write $\eta_t := \eta_t^{(i)}$ for ease of exposition. Then we have

$$\mathbb{E}_t(\eta_s - \eta_t) = \mathbb{E}_t\left(\int_t^s d\eta_u\right) = ab(s-t) - a\int_t^s \mathbb{E}_t(\eta_u)du.$$

Next, we define the random variable $Z_u := \mathbb{E}_t(\eta_u)$. Then, we have

$$dZ_s = a(b - Z_s)ds, \qquad Z_t = \eta_t.$$

Hence, we get

$$Z_s = \mathbb{E}_t(\eta_s) = Z_t e^{-a(s-t)} + b(1 - e^{-a(s-t)}) = \eta_t e^{-a(s-t)} + b(1 - e^{-a(s-t)}),$$

and

$$\mathbb{E}_t(\eta_s) - \eta_t = \eta_t(e^{-a(s-t)} - 1) + b(1 - e^{-a(s-t)}) = -a\epsilon_a(s-t)\eta_t + ab\epsilon_a(s-t).$$

Consequently, we obtain

$$\lambda\int_t^{t+h} \epsilon_\lambda(t+h-s)\left(\mathbb{E}(L_1)\mathbb{E}_t\left(\eta_s - \eta_t\right)\right)ds = \mathbb{E}(L_1)G(h)(b - \eta_t),$$

where

$$\begin{aligned} G(h) = G(h,a,\lambda) &= \lambda a\int_t^{t+h} \epsilon_\lambda(t+h-s)\epsilon_a(s-t)ds \\ &= -\frac{a+\lambda}{a\lambda} + h - \frac{a}{\lambda(\lambda-a)}e^{-\lambda h} + \frac{\lambda}{a(\lambda-a)}e^{-ah}. \end{aligned} \tag{20}$$

The results under the risk–neutral measure are essentially the same using the risk–neutral parameters $a^{\mathbb{Q}}, b^{\mathbb{Q}}$. Also, we denote by $G^{\mathbb{Q}}(h) := G(h, a^{\mathbb{Q}}, \lambda)$ the function G defined in (20) evaluated at the risk–neutral parameter. Altogether, we have

$$VRP_{t,t+h}^{\mathbb{Q},(i)c} = \eta_t^{(i)}F^{(1)}(h) + F^{(2)}(h),$$

where for $\kappa_1 = \mathbb{E}(L_1)$ and $\kappa_1^{\mathbb{Q}} = \mathbb{E}(L_1)$

$$\begin{aligned} F^{(1)}(h) &:= F^{(1)}(h,\lambda,a,a^{\mathbb{Q}},\kappa_1,\kappa_1^{\mathbb{Q}}) = \left(\kappa_1 - \kappa_1^{\mathbb{Q}}\right)\left(1 - \frac{\epsilon_\lambda(h)}{h}\right) - \frac{1}{h}\left(\kappa_1 G(h) - \kappa_1^{\mathbb{Q}}G^{\mathbb{Q}}(h)\right), \\ F^{(2)}(h) &:= F^{(2)}(h,\lambda,a,a^{\mathbb{Q}},b,b^{\mathbb{Q}},\kappa_1,\kappa_1^{\mathbb{Q}}) = \frac{1}{h}\left[\kappa_1 G(h)b - \kappa_1^{\mathbb{Q}}G^{\mathbb{Q}}(h)b^{\mathbb{Q}}\right]. \end{aligned}$$

□

Proof of Corollary 11 Throughout the proof we assume that $s \geq t$. Again, we skip the superscript and write $\eta_t := \eta_t^{(i)}$ for ease of exposition. Then we have

$$\mathbb{E}_t(\eta_s - \eta_t) = \mathbb{E}_t\left(\int_t^s d\eta_u\right) = -a\int_t^s \mathbb{E}_t(\eta_u)du + a(s-t)\mathbb{E}(L_1^\eta).$$

So, when we define $b := \mathbb{E}(L_1^\eta)$, we get exactly the same results as in the previous Corollary and we can define

$$G^{(1)}(h) := F^{(1)}(h), \qquad \text{and} \qquad G^{(2)}(h) := F^{(2)}(h).$$

□

References

Aït-Sahalia, Y. & Jacod, J. (2009), 'Testing for jumps in a discretely observed process', *Annals of Statistics* **37**(1), 184–222.

Bandi, F. M. & Renò, R. (2008), Nonparametric stochastic volatility. Working paper (June 2008). Available at SSRN: http://ssrn.com/abstract=1158438.

Barndorff-Nielsen, O. E. (2001), 'Superposition of Ornstein–Uhlenbeck type processes', *Theory of Probability and its Applications* **45**, 175–194.

Barndorff-Nielsen, O. E. (2010), Lévy bases and extended subordination. Thiele Centre Research Report No. 12, Aarhus University.

Barndorff-Nielsen, O. E. (2011), 'Stationary infinitely divisible processes', REBRAPE *Brazilian Journal of Probability and Statistics* **25**, 294–322.

Barndorff-Nielsen, O. E., Benth, F. E. & Veraart, A. E. D. (2010), Modelling energy spot prices by Lévy semistationary processes. CREATES Research Paper 2010–18. Available at SSRN: http://ssrn.com/abstract=1597700.

Barndorff-Nielsen, O. E., Benth, F. E. & Veraart, A. E. D. (2011*a*), Ambit processes and stochastic partial differential equations, *in* G. Di Nunno & B. Øksendal, eds, 'Advanced Mathematical Methods for Finance', Springer, pp. 35–74.

Barndorff-Nielsen, O. E., Benth, F. E. & Veraart, A. E. D. (2011*b*), Modelling electricity forward markets by ambit fields. Available at SSRN: http://ssrn.com/abstract=1938704.

Barndorff-Nielsen, O. E. & Leonenko, N. (2005), 'Spectral properties of superpositions of Ornstein–Uhlenbeck type processes', *Methodology and Computing in Applied Probability* **7**, 335–352.

Barndorff-Nielsen, O. E. & Pedersen, J. (2011), 'Meta–times and extended subordination', *Teoriya Veroyatnostei i ee Primeneniya* **56**, 398–407.

Barndorff-Nielsen, O. E. & Shephard, N. (2001), 'Non–Gaussian Ornstein–Uhlenbeck–based models and some of their uses in financial economics (with discussion)', *Journal of the Royal Statistical Society Series B* **63**, 167–241.

Barndorff-Nielsen, O. E. & Shephard, N. (2002), 'Econometric analysis of realised volatility and its use in estimating stochastic volatility models', *Journal of the Royal Statistical Society* **B 64**, 253–280.

Barndorff-Nielsen, O. E. & Shephard, N. (2003), 'Integrated OU processes', *Scandinavian Journal of Statistics* **30**, 277–295.

Barndorff-Nielsen, O. E. & Shephard, N. (2007), Variation, jumps, market frictions and high frequency data in financial econometrics, *in* R. Blundell, P. Torsten & W. K. Newey, eds, 'Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress', Econometric Society Monographs, Cambridge University Press, pp. 328–372.

Barndorff-Nielsen, O. E. & Shephard, N. (2012), *Financial Volatility in Continuous Time*, Cambridge University Press. To appear.

Barndorff-Nielsen, O. E. & Shiryaev, A. (2010), *Change of time and change of measure*, Advanced Series on Statistical Science & Applied Probability, 13, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ.

Barndorff-Nielsen, O. E. & Stelzer, R. (2011), 'Multivariate supOU processes', *Annals of Applied Probability* **21**, 140–182.

Barndorff-Nielsen, O. E. & Stelzer, R. (2012), 'The multivariate supOU stochastic volatility model', *Mathematical Finance* . Forthcoming.

Behme, A., Lindner, A. & Maller, R. (2011), 'Stationary solutions of the stochastic differential equation $dV_t = V_{t-}dU_t + dL_t$ with Lévy noise', *Stochastic Processes and their Applications* **121**, 91–108.

Benth, F. E. (2011), 'The stochastic volatility model of Barndorff–Nielsen and Shephard in commodity markets', *Mathematical Finance* **21**(4), 595–625.

Black, F. (1976), 'Studies of stock market volatility changes', *Proceedings of the Business and Economic Statistics Section, American Statistical Association* pp. 177–181.

Bollerslev, T., Litvinova, J. & Tauchen, G. (2006), 'Leverage and volatility feedback effects in high-frequency data', *Journal of Financial Econometrics* **4**(3), 353–384.

Bollerslev, T., Tauchen, G. & Zhou, H. (2009), 'Expected stock returns and variance risk premia', *Review of Financial Studies* **22**(11), 4463–4492.

Bouchaud, J.-P., Matacz, A. & Potters, M. (2001), 'Leverage effect in financial markets: The retarded volatility model', *Physical Review Letters* **87**(228701).

Carr, P. & Madan, D. (1999), 'Option valuation and the fast Fourier transform', *Journal of Computational Finance* **2**.

Carr, P. & Wu, L. (2009), 'Variance risk premiums', *Review of Financial Studies* **22**(3), 1311–1341.

Christensen, K., Oomen, R. & Podolskij, M. (2011), Fact or friction: Jumps at ultra high frequency. Preprint available at SSRN: http://ssrn.com/abstract=1848774.

Christie, A. A. (1982), 'The stochastic behavior of common stock variances', *Journal of Financial Economics* **10**, 407–432.

Christoffersen, P., Heston, S. & Jacobs, K. (2009), 'The shape and term structure of the index option smirk: why multifactor stochastic volatility models work so well', *Management Science* **55**(12), 1914–1932.

Christoffersen, P., Jacobs, K., Ornthanalai, C. & Wang, Y. (2008), 'Option valuation with long-run and short-run volatility components', *Journal of Financial Economics* **90**(3), 272 – 297.

Comte, F. & Renault, E. (1998), 'Long memory in continuous–time stochastic volatility models', *Mathematical Finance* **8**, 291–257.

Cox, J. (1996), 'The constant elasticity of variance option pricing model', *Journal of Portfolio Management* pp. 15–17.

Cox, J. C., Ingersoll, J. E. J. & Ross, S. A. (1985), ‘A theory of the term structure of interest rates’, *Econometrica* **53**(2), 385–408.

Drechsler, I. & Yaron, A. (2011), ‘What’s vol got to do with it’, *Review of Financial Studies* **24**(1), 1–45.

Duffie, D., Filipovic, D. & Schachermayer, W. (2003), ‘Affine processes and applications in finance’, *Annals of Applied Probability* **13**, 984–1053.

Dupire, B. (1994), ‘Pricing with a smile’, *Risk* **7**(1), 18–20.

Fasen, V. & Klüppelberg, C. (2007), Extremes of supOU processes, *in* F. Benth, G. Di Nunno, T. Lindström, B. Øksendal & T. Zhang, eds, ‘Stochastic Analysis and Applications. The Abel Symposium’, Vol. 2, Springer, Berlin, pp. 340–359.

Fasen, V., Klüppelberg, C. & Lindner, A. (2006), Extremal behavior of stochastic volatility models, *in* A. N. Shiryaev, M. R. Grossinho, P. E. Oliveira & M. L. Esquível, eds, ‘Stochastic Finance’, Springer US, pp. 107–155.

Ghysels, E., Harvey, A. & Renault, E. (1996), Stochastic volatility, *in* C. R. Maddala, G. S. and Rao, ed., ‘Handbook of Statistics’, Vol. 14 of *Statistical Methods in Finance*, North Holland, pp. 119–183.

Granger, C. J. (1981), ‘Some properties of time series data and their use in econometric model specification’, *Journal of Econometrics* **16**, 121–130.

Hagan, P., Kumar, D., Lesniewski, A. & Woodward, D. (2002), ‘Managing smile risk’, *Wilmott Magazine* pp. 84–108.

Harvey, A. & Shephard, N. (1996), ‘Estimation of an asymmetric stochastic volatility model for asset returns’, *Journal of Business and Economic Statistics* **14**, 429–434.

Heston, S. L. (1993), ‘A closed–form solution for options with stochastic volatility with applications to bond and currency options’, *Review of Financial Studies* **6**, 327–343.

Hubalek, F. & Sgarra, C. (2009), ‘On the Esscher transforms and other equivalent martingale measures for Barndorff–Nielsen and Shephard stochastic volatility models with jumps’, *Stochastic Processes and their Applications* **119**(7), 2137 – 2157.

Jacod, J. (1979), Changements de temps et changements d’espace, *in* ‘Calcul Stochastique et Problèmes de Martingales’, Vol. 714 of *Lecture Notes in Mathematics*, Springer.

Jacod, J. (2008), ‘Asymptotic properties of realized power variations and related functionals of semimartingales’, *Stochastic Processes and their Applications* **118**(4), 517–559.

Kallsen, J. (2006), A didactic note on affine stochastic volatility models, *in* Y. Kabanov, R. Liptser & J. Stoyanov, eds, ‘From Stochastic Calculus to Mathematical Finance’, Springer, Berlin, pp. 343–368.

Kallsen, J. & Shiraev, A. N. (2002), ‘Time change representation of stochastic integrals’, *Theory of Probability and its Applications* **46**(3), 522–528.

Karatzas, I. & Shreve, S. E. (1991), *Brownian Motion and Stochastic Calculus*, second edn, Springer, New York.

Keller-Ressel, M. (2008), Affine Processes — Theory and Applications in Finance, PhD thesis, Technische Universität Wien.

Klüppelberg, C. & Lindner, A. (2010), Stochastic volatility models: Extremal behavior, *in* 'Encyclopedia of Quantitative Finance', John Wiley & Sons, Ltd.

Kristensen, D. (2010), 'Nonparametric filtering of the realized spot volatility: A kernel–based approach', *Econometric Theory* **26**, 60–93.

Lee, R. W. (2004), 'Option pricing by transform methods: Extensions, unification, and error control', *Journal of Computational Finance* **7**, 51–86.

Lee, S. S. & Mykland, P. A. (2008), 'Jumps in financial markets: A new nonparametric test and jump dynamics', *Review of Financial Studies* **21**(6), 2535–2563.

Lewis, A. L. (2001), A simple option formula for general jump–diffusion and other exponential Lévy processes. Preprint available at SSRN: http://ssrn.com/abstract=282110 or doi:10.2139/ssrn.282110.

Meddahi, N. & Renault, E. (2004), 'Temporal aggregation of volatility models', *Journal of Econometrics* **119**, 355–379.

Muhle-Karbe, J., Pfaffel, O. & Stelzer, R. (2012), 'Option pricing in multivariate stochastic volatility models of OU type', *SIAM Journal on Financial Mathematics* . Forthcoming.

Nelson, D. B. (1991), 'Conditional heteroscedasticity in asset returns: a new approach', *Econometrica* **59**(2), 347–370.

Nicolato, E. & Venardos, E. (2003), 'Option pricing in stochastic volatility models of the Ornstein–Uhlenbeck type', *Mathematical Finance* **13**, 445–466.

Pedersen, J. (2003), The Lévy–Itô decomposition of an independently scattered random measure. MaPhySto preprint MPS–RR 2003–2.

Pigorsch, C. & Stelzer, R. (2009), A multivariate Ornstein–Uhlenbeck type stochastic volatility model. Preprint.

Raible, S. (2000), Lévy Processes in Finance: Theory, Numerics, and Empirical Facts, PhD thesis, University of Freiburg.

Rajput, B. & Rosinski, J. (1989), 'Spectral representation of infinitely divisible distributions', *Probability Theory and Related Fields* **82**, 451–487.

Sato, K. (1999), *Lévy processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge.

Shephard, N. & Andersen, T. G. (2009), Stochastic volatility: Origins and overview, *in* T. G. Andersen, R. Davis, J.-P. Kreiss & T. Mikosch, eds, 'Handbook of Financial Time Series', Springer, pp. 233–254.

Shephard, N., ed. (2005), *Stochastic Volatility: Selected Readings*, Oxford University Press, Oxford.

Stein, E. M. & Stein, J. C. (1991), 'Stock price distributions with stochastic volatility: An analytic approach', *Review of Financial Studies* **4**, 727–752.

Tauchen, G. (2004), Remarks on recent developments in stochastic volatility: statistical modelling and general equilibrium. Working paper, Department of Economics, Duke University.

Tauchen, G. (2005), Stochastic volatility and general equilibrium. Working paper, Department of Economics, Duke University.

Todorov, V. (2010), 'Variance risk premium dynamics: The role of jumps', *The Review of Financial Studies* **23**(1), 345–383.

Veraart, A. E. D. (2010), 'Inference for the jump part of quadratic variation of Itô semimartingales', *Econometric Theory* **26**, 331 –368.

Veraart, A. E. D. & Veraart, L. A. M. (2012), 'Stochastic volatility and stochastic leverage', *Annals of Finance* . Forthcoming.

Veraart, A. & Winkel, M. (2010), Time change, *in* R. Cont, ed., 'Encyclopedia of Quantitative Finance', Vol. 4, Wiley, pp. 1812–1816.

Walsh, J. (1986), An introduction to stochastic partial differential equations, *in* R. Carmona, H. Kesten & J. Walsh, eds, 'Lecture Notes in Mathematics 1180', Ecole d'Eté de Probabilités de Saint–Flour XIV — 1984, Springer.

Wu, L. (2011), 'Variance dynamics: Joint evidence from options and high–frequency returns', *Journal of Econometrics* **160**, 280–287.

Yu, J. (2005), 'On leverage in a stochastic volatility model', *Journal of Econometrics* **127**(2), 165–178.